ACTIVATING MUTATIONS OF PLATELET DERIVED GROWTH FACTOR RECEPTOR ALPHA (PDGFRA) AS DIAGNOSTIC MARKERS AND THERAPEUTIC TARGETS

CROSS REFERENCE TO RELATED APPLICATIONS

This application claims the benefit of U.S. Provisional Applications No. 60/389,107, filed July 13, 2002, and No. 60/438,899, filed January 8, 2003. These provisional applications are incorporated herein in their entirety.

STATEMENT OF GOVERNMENT SUPPORT

This invention was made with United States government support pursuant to employment of one of the inventors as a Federal employee, as well as grant funding from a Veterans Affairs Merit Review Grant; the United States government has certain rights in the invention.

FIELD

This disclosure relates to tyrosine kinases, particularly receptor tyrosine kinases with one or more activation mutations. Further, it relates to methods of using these molecules in screens and analyses, including diagnoses, prognoses, and systems for identification and/or selection of pharmaceutical compounds.

BACKGROUND OF THE DISCLOSURE

Tyrosine kinases are expressed by many human cancers. These enzymes are attractive targets for the development of anticancer drugs, as it has been possible to optimize compounds with excellent inhibitory potency and selectivity to individual target tyrosine kinases. The utility of this approach has been highlighted by the success of imatinib mesylate (Gleevec™) in the treatment of chronic myelogenous leukemia (CML) and gastrointestinal stromal tumors (GISTs).

Expression of tyrosine kinases is ubiquitous in both cancers and normal tissues. Therefore, the efficacy of a kinase inhibitor is dependent on two factors: 1) the degree to which the target kinase is activated in a particular cancer, and 2) the degree to which the growth and survival of the cancer cells is dependent on the activated target kinase.

Gastrointestinal stromal tumors provide an excellent example of this principle. KIT tyrosine kinase is detectable by immunohistochemistry in a wide variety of cancers and normal tissues, but mutations of the KIT gene that yield constitutively active KIT kinase are found in only a small subset of tumors (Heinrich *et al.*, *J. Clin. Oncol.*, 20: 1692-1703, 2002). More than 85% of GISTs harbor such activating mutations (Blanke *et al.*, *Proceedings of ASCO* 20, 1a-1a. 2001; Heinrich *et al.*, *J. Clin. Oncol.*, 20: 1692-1703, 2002; Hirota *et al.*, *J. Pathol.*, 193: 505-510, 2001; Rubin *et al.*, *Cancer Res*, 61: 8118-8121, 2001) and, correspondingly, phosphorylation of KIT kinase (a marker of activation) was recently demonstrated in most fresh-frozen GIST specimens (Rubin *et al.*, *Cancer*

Res, 61: 8118-8121, 2001). Such phosphorylation of KIT is rarely observed in other cancer specimens. Recent success in the treatment of advanced malignant GISTs with imatinib mesylate is thought to reflect an important role of KIT activation in the growth and/or survival of GIST tumor cells (Blanke *et al.*, *Proceedings of ASCO* 20, 1a-1a. 2001; Joensuu *et al.*, *N Engl J Med*, 1052: 1052-1056, 2001; Van Oosterom *et al.*, *Lancet*, 358:1421-1423, 2001). The observation that treatment results with imatinib mesylate are significantly better for tumors with evidence of mutational activation of KIT than for tumors with no KIT mutation further supports this view (Heinrich *et al.*, *J. Clin. Oncol.*, 20: 1692-1703, 2002). Thus, in the case of GISTs, testing of clinical specimens for genomic mutations resulting in tyrosine kinase activation will be useful in determining which patients are most likely to respond to a tyrosine kinase inhibitor.

The PDGFRA (or PDGFR-α) protein is a type III receptor tyrosine kinase with homology to KIT, FLT3, CSF1-R and PDGFR-β (PDGFRB). Although PDGFRA activation has been hypothesized to be involved in certain cancers, most notably gliomas, evidence of genomic activation in human cancer has only recently been reported in two cases of myeloproliferative disease associated with translocation of the BCR and PDGFRA genes.

SUMMARY OF THE DISCLOSURE

Disclosed herein are novel mutations of PDGFRA that result in constitutive activation of this tyrosine kinase. These mutations were initially discovered in GISTs. Also disclosed are consensus PDGFRA nucleic acid and amino acid sequences, which summarize certain groups of activating mutations and regions of relatively active mutation.

Thus, this disclosure provides several novel PDGFRA variant proteins, and nucleic acids encoding these variants. Also disclosed are methods of using these molecules in detecting biological conditions associated with an activating PDGFRA mutation in a subject, methods of treating such conditions, methods of selecting treatments (*e.g.*, specific tyrosine kinase inhibitors), and methods of screening for inhibitors of PDGFRA activity, particularly activated PDGFRA variant activity. Oligonucleotides for use in examples of such methods are also provided.

Also disclosed herein are protein specific binding agents, such as antibodies, that bind specifically to at least one epitope of a PDGFRA variant protein preferentially compared to wildtype PDGFRA, and methods of using such antibodies in diagnosis, treatment, and screening.

Kits are also provided for carrying out the methods described herein.

The foregoing and other features and advantages will become more apparent from the following detailed description of several embodiments, which proceeds with reference to the accompanying figures.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1: **Immunostaining for phosphotyrosine (A) and PDGFRA (B) in GIST478.** A) A strongly tyrosine phosphorylated doublet at 150/170 kD is seen in the RTK immunoprecipitate

(lane 2). This phosphorylated doublet corresponds to two of the stronger phosphoproteins in the total cell lysate (lane 1), and comigrates with the strongly phosphorylated PDGFRA doublet (lane 4). KIT is not demonstrably phosphorylated (lane 3). **B)** The strongly phosphorylated RTK (lane 2) was confirmed as PDGFRA, by stripping and restaining the blot with a specific antibody to PDGFRA.

Figure 2: **Sequential staining of GIST immunoblot for KIT (A), phosphoPDGFRA Y754 (B), and total PDGFRA (C). A)** The four GISTs analyzed here include two cases with a low (lane 1) or absent (lane 2) level of KIT expression and two cases with strong KIT expression (lanes 3 and 4). **B)** Strongly phosphorylated PDGFRA (doublet at 150/170 kD) is seen in the GISTs with low-to-absent KIT expression. **C)** Total PDGFRA is also expressed strongly in the two GISTs with low-to-absent KIT expression. The two GISTs with phosphoPDGFRA have D842V oncogenic mutations.

Figure 3**: Detection of PDGFRA activation loop deletion mutations by D-HPLC.** DNA was isolated from GISTs and amplified using primer pair PDGFRA 181634F and PDGFRA 181874R as described herein. Amplicons were analyzed at 50 °C using a Transgenomics WAVE™ D-HPLC system. Sample 1 has the DIMH deletion described herein. The deletion mutant is readily detected due to the appearance of novel peaks representing species homozygous for the deletion and heteroduplexes of wild-type and deletion mutation.

Figure 4**: Detection of PDGFRA activation loop V824V SNP and D842V point mutation by D-HPLC.** Amplicons were prepared from GISTs using the PDGFRA 181634F and PDGFRA 18174R primer pair as described above and analyzed at 61 °C using a Transgenomics WAVE™ D-HPLC system. Under partially denaturing conditions, amplicons with the V824V SNP and the D842V point mutation (two examples) elute in a complex pattern. The V824V and D842V amplicons have unique elution profiles. Direct DNA sequencing was performed to confirm that the V824V and D842V amplicons contained the equivalent stretch of PDGFRA nucleotide sequence.

Figure 5**: Detection of D842V point mutation using a primer pair that excludes the V824V SNP.** Amplicons were prepared from GISTs using the PDGFRA 181752F (SNP exclusion) and PDGFRA 181874R primer pair as described above and analyzed at 61°C using a Transgenomics WAVE™ D-HPLC system. Under partially denaturing conditions, amplicons with the D842V point mutation elute in a complex pattern. Note that this amplicon does not contain the V824V SNP and therefore these amplicons have the same elution profile as for wild-type PDGFRA. Direct DNA sequencing was performed to confirm that the amplicons from GISTs with V824V (two examples) versus D842V contained the equivalent stretch of PDGFRA nucleotide sequence.

Figure 6**: Detection of PDGFRA Exon 12 Deletion and Insertion Variants.** Amplicons were prepared from GISTs using the PDGFRA 170636F and PDGFRA 170894R primer pair as described above and analyzed at 50 °C using a Transgenomics WAVE™ D-HPLC system. The amplicons prepared from the two samples with wild-type PDGFRA exon 12 elute as a single peak. In contrast, amplicons from tumors with either a deletion mutation or an insertion are easily detected due to the appearance of novel peaks representing species homozygous for the deletion and

heteroduplexes of wild-type and deletion mutation. In tumors homozygous for these mutations, only a single unique elution peak would be detected. These mutations would identifiable based on the unique peak elution profile compared with wild type amplicons.

Figure 7: **Differential sensitivity of various KIT activation loop mutants to imatinib mesylate.** ***Figure 7*** shows the genomic sequences of PDGFRA around exon 18 (FIG 7A) and exon 12 (FIG 7B). PDGFRA primers are indicated; PDGFRA exon sequences and amino acid translations are also shown.

Figure 8**: DGFRA mutations in GISTs result in constitutive activation of PDGFRA kinase.** ***Figure 8*** shows a series of immunoblots, probed with antibodies to phosphor-tyrosine and PDGFRA. CHO cells were transiently transfected with expression vectors encoding cDNAs for wild-type or mutant PDGFRA. Transfected cells were serum starved overnight and treated with vehicle or ligand (recombinant human PDGF-AA) for 10 minutes. Whole cell lysates were immunostained sequentially for phospho-tyrosine and PDGFRA. Wild type PDGFRA displays low-level phosphorylation that is upregulated by ligand stimulation with PDGF-AA. In contrast, the mutant PDGFRA proteins display ligand-independent phosphorylation.

Figure 9**: Cell signaling profiles in PDGFRA-mutant (2686, 478, and 1015) and KIT-mutant GISTs (174 and 208).** ***Figure 9*** shows a series of immunoblots, illustrating the cell signaling profiles of the indicated mutants. Whole cell lysates were prepared from snap-frozen GISTs, and immunoblots were detected with antibodies to phosphorylated and total forms of AKT, MAPK, and STATs. All GISTs express phosphorylated AKT, MAPK, STAT1, and STAT3, whereas STAT5 is not tyrosine phosphorylated.

SEQUENCE LISTING

The nucleic and amino acid sequences listed in the accompanying sequence listing are shown using standard letter abbreviations for nucleotide bases, and three letter code for amino acids, as defined in 37 C.F.R. 1.822. Only one strand of each nucleic acid sequence is shown, but the complementary strand is understood as included by any reference to the displayed strand. Unless specifically noted otherwise herein, the position numbering associated with the name of a variant PDGFRA molecule is based on numbering in the corresponding wildtype molecule. Where a reference is made to positions in a variant, the numbering is based on the actual position in the specified variant. In the accompanying sequence listing:

SEQ ID NO: 1 shows the nucleic acid sequence of the human PDGFRA cDNA (GenBank Accession No. XM_011186); the sequence list also shows the encoded protein.

SEQ ID NO: 2 shows the amino acid sequence of human PDGFRA protein.

SEQ ID NO: 3 shows the nucleic acid sequence of the human PDGFRA D842V variant cDNA; the sequence list also shows the encoded protein.

SEQ ID NO: 4 shows the amino acid sequence of human PDGFRA D842V variant protein.

SEQ ID NO: 5 shows the nucleic acid sequence of the human PDGFRA DIMH842-845 variant cDNA; the sequence list also shows the encoded protein.

SEQ ID NO: 6 shows the amino acid sequence of human PDGFRA DIMH842-845 variant protein.

SEQ ID NO: 7 shows the nucleic acid sequence of the human PDGFRA HSDN845-848P variant cDNA; the sequence list also shows the encoded protein.

SEQ ID NO: 8 shows the amino acid sequence of human PDGFRA HSDN845-848P variant protein.

SEQ ID NO: 9 shows the nucleic acid sequence of the human PDGFRA ER561-562 variant cDNA; the sequence list also shows the encoded protein.

SEQ ID NO: 10 shows the amino acid sequence of human PDGFRA ER561-562 variant protein.

SEQ ID NO: 11 shows the nucleic acid sequence of the human PDGFRA SPDGHE566-571R variant cDNA; the sequence list also shows the encoded protein.

SEQ ID NO: 12 shows the amino acid sequence of human PDGFRA SPDGHE566-571R variant protein.

SEQ ID NOs: 13-18 are amino acid sequences of the RTK catalytic domain sequences of different families of human RTK proteins.

SEQ ID NO: 19 is the genomic sequence of PDGFRA, with introns and exons indicated. Regions where the sequence is unknown or unconfirmed have been indicated with "n" designations using standard conventions. This sequence is available in the April 2002 release of the human genome project, as provided by University of California, Santa Cruz, on their Internet website.

SEQ ID NO: 20 shows the nucleic acid sequence of the human PDGFRA V561D variant cDNA; the sequence list also shows the encoded protein.

SEQ ID NO: 21 shows the amino acid sequence of human PDGFRA V561D variant protein.

SEQ ID NO: 22 shows the nucleic acid sequence of the human PDGFRA RVIES560-564 variant cDNA; the sequence list also shows the encoded protein.

SEQ ID NO: 23 shows the amino acid sequence of human PDGFRA RVIES560-564 variant protein.

SEQ ID NO: 24 shows the nucleic acid sequence of the human PDGFRA Deletion RD841-842KI variant cDNA; the sequence list also shows the encoded protein.

SEQ ID NO: 25 shows the amino acid sequence of human PDGFRA Deletion RD841-842KI variant protein.

SEQ ID NO: 26 shows the consensus sequence produced by aligning the nucleic acid sequences of each of the identified activating PDGFRA mutants (SEQ ID NOs: 3, 5, 7, 9, 11, 20, 22, and 24), and the consensus protein encoded thereby.

SEQ ID NO: 27 shows a PDGFRA consensus sequence.

DETAILED DESCRIPTION

I. Abbreviations

2D-PAGE	two-dimensional polyacrylamide gel electrophoresis
ASO	allele-specific oligonucleotide
ASOH	allele-specific oligonucleotide hybridization
DASH	dynamic allele-specific hybridization
ELISA	enzyme-linked immunosorbant assay
HPLC	high pressure liquid chromatography
MALDI-TOF	matrix-assisted laser desorption/ionization time-of-flight
PCR	polymerase chain reaction
PDGFRA	platelet derived growth factor receptor alpha
PDGFRB	platelet derived growth factor receptor beta
RT-PCR	reverse-transcription polymerase chain reaction
SSCP	single-strand conformation polymorphism
TKI	tyrosine kinase inhibitor

II. Terms

Unless otherwise noted, technical terms are used according to conventional usage. Definitions of common terms in molecular biology may be found in Benjamin Lewin, *Genes V*, published by Oxford University Press, 1994 (ISBN 0-19-854287-9); Kendrew *et al.* (eds.), *The Encyclopedia of Molecular Biology*, published by Blackwell Science Ltd., 1994 (ISBN 0-632-02182-9); and Robert A. Meyers (ed.), *Molecular Biology and Biotechnology: a Comprehensive Desk Reference*, published by VCH Publishers, Inc., 1995 (ISBN 1-56081-569-8).

In order to facilitate review of the various embodiments of the invention, the following explanations of specific terms are provided:

Antisense, Sense, and Antigene: Double-stranded DNA (dsDNA) has two strands, a 5' -> 3' strand, referred to as the plus strand, and a 3' -> 5' strand (the reverse complement), referred to as the minus strand. Because RNA polymerase adds nucleic acids in a 5' -> 3' direction, the minus strand of the DNA serves as the template for the RNA during transcription. Thus, the RNA formed will have a sequence complementary to the minus strand and identical to the plus strand (except that U is substituted for T).

Antisense molecules are molecules that are specifically hybridizable or specifically complementary to either RNA or the plus strand of DNA. Sense molecules are molecules that are specifically hybridizable or specifically complementary to the minus strand of DNA. Antigene molecules are either antisense or sense molecules directed to a dsDNA target.

cDNA (complementary DNA): A piece of DNA lacking internal, non-coding segments (introns) and transcriptional regulatory sequences. cDNA may also contain untranslated regions (UTRs) that are responsible for translational control in the corresponding RNA molecule. cDNA is usually synthesized in the laboratory by reverse transcription from messenger RNA extracted from cells.

DNA (deoxyribonucleic acid): DNA is a long chain polymer which comprises the genetic material of most living organisms (some viruses have genes comprising ribonucleic acid (RNA)).

The repeating units in DNA polymers are four different nucleotides, each of which comprises one of the four bases, adenine (A), guanine (G), cytosine (C), and thymine (T) bound to a deoxyribose sugar to which a phosphate group is attached. Triplets of nucleotides (referred to as codons) code for each amino acid in a polypeptide, or for a stop signal. The term codon is also used for the corresponding (and complementary) sequences of three nucleotides in the mRNA into which the DNA sequence is transcribed.

Unless otherwise specified, any reference to a DNA molecule is intended to include the reverse complement of that DNA molecule. Except where single-strandedness is required by the text herein, DNA molecules, though written to depict only a single strand, encompass both strands of a double-stranded DNA molecule. Thus, a reference to the nucleic acid molecule that encodes a specific protein, or a fragment thereof, encompasses both the sense strand and its reverse complement. For instance, it is appropriate to generate probes or primers from the reverse complement sequence of the disclosed nucleic acid molecules.

Hybridization: Oligonucleotides and their analogs hybridize by hydrogen bonding, which includes Watson-Crick, Hoogsteen or reversed Hoogsteen hydrogen bonding, between complementary bases. Generally, nucleic acid consists of nitrogenous bases that are either pyrimidines (cytosine (C), uracil (U), and thymine (T)) or purines (adenine (A) and guanine (G)). These nitrogenous bases form hydrogen bonds between a pyrimidine and a purine, and the bonding of the pyrimidine to the purine is referred to as "base pairing." More specifically, A will hydrogen bond to T or U, and G will bond to C. "Complementary" refers to the base pairing that occurs between to distinct nucleic acid sequences or two distinct regions of the same nucleic acid sequence.

"Specifically hybridizable" and "specifically complementary" are terms that indicate a sufficient degree of complementarity such that stable and specific binding occurs between the oligonucleotide (or its analog) and the DNA or RNA target. The oligonucleotide or oligonucleotide analog need not be 100% complementary to its target sequence to be specifically hybridizable. An oligonucleotide or analog is specifically hybridizable when binding of the oligonucleotide or analog to the target DNA or RNA molecule interferes with the normal function of the target DNA or RNA, and there is a sufficient degree of complementarity to avoid non-specific binding of the oligonucleotide or analog to non-target sequences under conditions where specific binding is desired, for example under physiological conditions in the case of *in vivo* assays or systems. Such binding is referred to as specific hybridization.

Hybridization conditions resulting in particular degrees of stringency will vary depending upon the nature of the hybridization method of choice and the composition and length of the hybridizing nucleic acid sequences. Generally, the temperature of hybridization and the ionic strength (especially the Na^+ concentration) of the hybridization buffer will determine the stringency of hybridization, though waste times also influence stringency. Calculations regarding hybridization conditions required for attaining particular degrees of stringency are discussed by Sambrook *et al.*

(ed.), *Molecular Cloning: A Laboratory Manual*, 2nd ed., vol. 1-3, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989, chapters 9 and 11, herein incorporated by reference.

For present purposes, "stringent conditions" encompass conditions under which hybridization will only occur if there is less than 25% mismatch between the hybridization molecule and the target sequence. "Stringent conditions" may be broken down into particular levels of stringency for more precise definition. Thus, as used herein, "moderate stringency" conditions are those under which molecules with more than 25% sequence mismatch will not hybridize; conditions of "medium stringency" are those under which molecules with more than 15% mismatch will not hybridize, and conditions of "high stringency" are those under which sequences with more than 10% mismatch will not hybridize. Conditions of "very high stringency" are those under which sequences with more than 6% mismatch will not hybridize.

Injectable composition: A pharmaceutically acceptable fluid composition including at least one active ingredient. The active ingredient is usually dissolved or suspended in a physiologically acceptable carrier, and the composition can additionally include amounts of one or more non-toxic auxiliary substances, such as emulsifying agents, preservatives, and pH buffering agents and the like. Such injectable compositions that are useful for use with the provided nucleotides and proteins are conventional; appropriate formulations are well known in the art.

***In vitro* amplification:** Techniques that increase the number of copies of a nucleic acid molecule in a sample or specimen. An example of *in vitro* amplification is the polymerase chain reaction, in which a biological sample collected from a subject is contacted with a pair of oligonucleotide primers, under conditions that allow for the hybridization of the primers to nucleic acid template in the sample. The primers are extended under suitable conditions, dissociated from the template, and then re-annealed, extended, and dissociated to amplify the number of copies of the nucleic acid.

The product of *in vitro* amplification may be characterized by electrophoresis, restriction endonuclease cleavage patterns, oligonucleotide hybridization or ligation, and/or nucleic acid sequencing, using standard techniques.

Other examples of *in vitro* amplification techniques include strand displacement amplification (see U.S. Patent No. 5,744,311); transcription-free isothermal amplification (see U.S. Patent No. 6,033,881); repair chain reaction amplification (see WO 90/01069); ligase chain reaction amplification (see EP-A-320 308); gap filling ligase chain reaction amplification (see U.S. Patent No. 5,427,930); coupled ligase detection and PCR (see U.S. Patent No. 6,027,889); and NASBA™ RNA transcription-free amplification (see U.S. Patent No. 6,025,134).

Isolated: An "isolated" biological component (such as a nucleic acid molecule, protein or organelle) has been substantially separated or purified away from other biological components in the cell of the organism in which the component naturally occurs, *i.e.*, other chromosomal and extra-chromosomal DNA and RNA, proteins and organelles. Nucleic acids and proteins that have been "isolated" include nucleic acids and proteins purified by standard purification methods. The term

also embraces nucleic acids and proteins prepared by recombinant expression in a host cell as well as chemically synthesized nucleic acids.

Mutation: Any change of the DNA sequence within a gene or chromosome. In some instances, a mutation will alter a characteristic or trait (phenotype), but this is not always the case. Types of mutations include base substitution point mutations (*e.g.*, transitions or transversions), deletions, and insertions. Missense mutations are those that introduce a different amino acid into the sequence of the encoded protein; nonsense mutations are those that introduce a new stop codon. In the case of insertions or deletions, mutations can be in-frame (not changing the frame of the overall sequence) or frame shift mutations, which may result in the misreading of a large number of codons (and often leads to abnormal termination of the encoded product due to the presence of a stop codon in the alternative frame).

This term specifically encompasses variations that arise through somatic mutation, for instance those that are found only in disease cells, but not constitutionally, in a given individual. Examples of such somatically-acquired variations include the point mutations that frequently result in altered function of various genes that are involved in development of cancers. This term also encompasses DNA alterations that are present constitutionally, that alter the function of the encoded protein in a readily demonstrable manner, and that can be inherited by the children of an affected individual. In this respect, the term overlaps with "polymorphism," as defined below, but generally refers to the subset of constitutional alterations that have arisen within the past few generations in a kindred and that are not widely disseminated in a population group. In particular embodiments, the term is directed to those constitutional alterations that have major impact on the health of affected individuals.

Nucleotide: "Nucleotide" includes, but is not limited to, a monomer that includes a base linked to a sugar, such as a pyrimidine, purine or synthetic analogs thereof, or a base linked to an amino acid, as in a peptide nucleic acid (PNA). A nucleotide is one monomer in a polynucleotide. A nucleotide sequence refers to the sequence of bases in a polynucleotide.

Oligonucleotide: An oligonucleotide is a plurality of joined nucleotides joined by native phosphodiester bonds, between about 6 and about 500 nucleotides in length. An oligonucleotide analog refers to moieties that function similarly to oligonucleotides but have non-naturally occurring portions. For example, oligonucleotide analogs can contain non-naturally occurring portions, such as altered sugar moieties or inter-sugar linkages, such as a phosphorothioate oligodeoxynucleotide. Functional analogs of naturally occurring polynucleotides can bind to RNA or DNA, and include PNA molecules.

Particular oligonucleotides and oligonucleotide analogs can include linear sequences up to about 300 nucleotides in length, for example a sequence (such as DNA or RNA) that is at least 6 bases, for example at least 8, 10, 15, 20, 25, 30, 35, 40, 45, 50, 100 or even 200 or more bases long, or from about 6 to about 50 bases, for example about 10-25 bases, such as 12, 15, 20, or 25 bases.

Operably linked: A first nucleic acid sequence is operably linked with a second nucleic acid sequence when the first nucleic acid sequence is placed in a functional relationship with the second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Generally, operably linked DNA sequences are contiguous and, where necessary to join two protein-coding regions, in the same reading frame.

Open reading frame (ORF): A series of nucleotide triplets (codons) coding for amino acids without any internal termination codons. These sequences are usually translatable into a peptide.

Ortholog: Two nucleic acid or amino acid sequences are orthologs of each other if they share a common ancestral sequence and diverged when a species carrying that ancestral sequence split into two species. Orthologous sequences are also homologous sequences.

Parenteral: Administered outside of the intestine, *e.g.*, not via the alimentary tract. Generally, parenteral formulations are those that will be administered through any possible mode except ingestion. This term especially refers to injections, whether administered intravenously, intrathecally, intramuscularly, intraperitoneally, or subcutaneously, and various surface applications including intranasal, intradermal, and topical application, for instance.

Peptide Nucleic Acid (PNA): An oligonucleotide analog with a backbone comprised of monomers coupled by amide (peptide) bonds, such as amino acid monomers joined by peptide bonds.

Pharmaceutically acceptable carriers: The pharmaceutically acceptable carriers useful with the compositions provided herein are conventional. By way of example, Martin, in *Remington's Pharmaceutical Sciences*, published by Mack Publishing Co., Easton, PA, 19th Edition, 1995, describes compositions and formulations suitable for pharmaceutical delivery of the nucleotides and proteins herein disclosed.

In general, the nature of the carrier will depend on the particular mode of administration being employed. For instance, parenteral formulations usually comprise injectable fluids that include pharmaceutically and physiologically acceptable fluids such as water, physiological saline, balanced salt solutions, aqueous dextrose, glycerol or the like as a vehicle. For solid compositions (*e.g.*, powder, pill, tablet, or capsule forms), conventional non-toxic solid carriers can include, for example, pharmaceutical grades of mannitol, lactose, starch, or magnesium stearate. In addition to biologically-neutral carriers, pharmaceutical compositions to be administered can contain minor amounts of non-toxic auxiliary substances, such as wetting or emulsifying agents, preservatives, and pH buffering agents and the like, for example sodium acetate or sorbitan monolaurate.

Polymorphism: Variant in a sequence of a gene, usually carried from one generation to another in a population. Polymorphisms can be those variations (nucleotide sequence differences) that, while having a different nucleotide sequence, produce functionally equivalent gene products, such as those variations generally found between individuals, different ethnic groups, geographic locations. The term polymorphism also encompasses variations that produce gene products with

altered function, *i.e.*, variants in the gene sequence that lead to gene products that are not functionally equivalent. This term also encompasses variations that produce no gene product, an inactive gene product, or increased or increased activity gene product.

Polymorphisms can be referred to, for instance, by the nucleotide position at which the variation exists, by the change in amino acid sequence caused by the nucleotide variation, or by a change in some other characteristic of the nucleic acid molecule or protein that is linked to the variation (*e.g.*, an alteration of a secondary structure such as a stem-loop, or an alteration of the binding affinity of the nucleic acid for associated molecules, such as polymerases, RNases, and so forth).

Probes and primers: Nucleic acid probes and primers can be readily prepared based on the nucleic acid molecules provided as indicators of disease or disease progression. It is also appropriate to generate probes and primers based on fragments or portions of these nucleic acid molecules. Also appropriate are probes and primers specific for the reverse complement of these sequences, as well as probes and primers to 5' or 3' regions.

A probe comprises an isolated nucleic acid attached to a detectable label or other reporter molecule. Typical labels include radioactive isotopes, enzyme substrates, co-factors, ligands, chemiluminescent or fluorescent agents, haptens, and enzymes. Methods for labeling and guidance in the choice of labels appropriate for various purposes are discussed, *e.g.*, in Sambrook *et al.* (In *Molecular Cloning: A Laboratory Manual*, CSHL, New York, 1989) and Ausubel *et al.* (In *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, 1998).

Primers are short nucleic acid molecules, for instance DNA oligonucleotides 10 nucleotides or more in length. Longer DNA oligonucleotides may be about 15, 20, 25, 30 or 50 nucleotides or more in length. Primers can be annealed to a complementary target DNA strand by nucleic acid hybridization to form a hybrid between the primer and the target DNA strand, and then the primer extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification of a nucleic acid sequence, *e.g.*, by the polymerase chain reaction (PCR) or other *in vitro* nucleic-acid amplification methods known in the art.

Methods for preparing and using nucleic acid probes and primers are described, for example, in Sambrook *et al.* (In *Molecular Cloning: A Laboratory Manual*, CSHL, New York, 1989), Ausubel *et al.* (ed.) (In *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, 1998), and Innis *et al.* (*PCR Protocols, A Guide to Methods and Applications*, Academic Press, Inc., San Diego, CA, 1990). Amplification primer pairs (for instance, for use with polymerase chain reaction amplification) can be derived from a known sequence such as the PDGFRA or other tyrosine kinase sequences described herein, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, © 1991, Whitehead Institute for Biomedical Research, Cambridge, MA).

One of ordinary skill in the art will appreciate that the specificity of a particular probe or primer increases with its length. Thus, for example, a primer comprising 30 consecutive nucleotides of a tyrosine kinase protein encoding nucleotide will anneal to a target sequence, such as another

homolog of the designated tyrosine kinase protein, with a higher specificity than a corresponding primer of only 15 nucleotides. Thus, in order to obtain greater specificity, probes and primers can be selected that comprise at least 20, 23, 25, 30, 35, 40, 45, 50 or more consecutive nucleotides of a tyrosine kinase-encoding nucleotide sequences.

Also provided are isolated nucleic acid molecules that comprise specified lengths of tyrosine kinase-encoding nucleotide sequences. Such molecules may comprise at least 10, 15, 20, 23, 25, 30, 35, 40, 45 or 50 or more (*e.g.*, at least 100, 150, 200, 250, 300 and so forth) consecutive nucleotides of these sequences or more. These molecules may be obtained from any region of the disclosed sequences (*e.g.*, a PDGFRA nucleic acid may be apportioned into halves or quarters based on sequence length, and isolated nucleic acid molecules may be derived from the first or second halves of the molecules, or any of the four quarters, etc.). A cDNA or other encoding sequence also can be divided into smaller regions, *e.g.* about eighths, sixteenths, twentieths, fiftieths, and so forth, with similar effect.

Another mode of division, provided by way of example, is to divide a tyrosine kinase-encoding sequence based on the regions of the sequence that are relatively more or less homologous to other tyrosine kinase sequences.

Another mode of division is to select the 5' (upstream) and/or 3' (downstream) region associated with a tyrosine kinase gene (*e.g.*, PDGFRA).

Nucleic acid molecules may be selected that comprise at least 10, 15, 20, 25, 30, 35, 40, 50, 100, 150, 200, 250, 300 or more consecutive nucleotides of any of these or other portions of a PDGFRA nucleic acid molecule, such as those disclosed herein, and associated flanking regions. Thus, representative nucleic acid molecules might comprise at least 10 consecutive nucleotides of the PDGFRA cDNA shown in SEQ ID NO: 1.

Protein: A biological molecule expressed by a gene or recombinant or synthetic coding sequence and comprised of amino acids.

Purified: The term "purified" does not require absolute purity; rather, it is intended as a relative term. Thus, for example, a purified protein preparation is one in which the protein referred to is more pure than the protein in its natural environment within a cell or within a production reaction chamber (as appropriate).

Recombinant: A recombinant nucleic acid is one that has a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two otherwise separated segments of sequence. This artificial combination can be accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, *e.g.*, by genetic engineering techniques.

Sequence identity: The similarity between two nucleic acid sequences, or two amino acid sequences, is expressed in terms of the similarity between the sequences, otherwise referred to as sequence identity. Sequence identity is frequently measured in terms of percentage identity (or similarity or homology); the higher the percentage, the more similar the two sequences are.

Homologs or orthologs of human PDGFRA protein, and the corresponding cDNA or gene sequence(s), will possess a relatively high degree of sequence identity when aligned using standard methods. This homology will be more significant when the orthologous proteins or genes or cDNAs are derived from species that are more closely related (*e.g.*, human and chimpanzee sequences), compared to species more distantly related (*e.g.*, human and *C. elegans* sequences).

Methods of alignment of sequences for comparison are well known in the art. Various programs and alignment algorithms are described in: Smith & Waterman *Adv. Appl. Math.* 2: 482, 1981; Needleman & Wunsch *J. Mol. Biol.* 48: 443, 1970; Pearson & Lipman *Proc. Natl. Acad. Sci. USA* 85: 2444, 1988; Higgins & Sharp *Gene,* 73: 237-244, 1988; Higgins & Sharp *CABIOS* 5: 151-153, 1989; Corpet *et al. Nuc. Acids Res.* 16, 10881-90, 1988; Huang *et al. Computer Appls. in the Biosciences* 8, 155-65, 1992; and Pearson *et al. Meth. Mol. Bio.* 24, 307-31, 1994. Altschul *et al.* (*J. Mol. Biol.* 215:403-410, 1990), presents a detailed consideration of sequence alignment methods and homology calculations.

The NCBI Basic Local Alignment Search Tool (BLAST) (Altschul *et al. J. Mol. Biol.* 215:403-410, 1990) is available from several sources, including the National Center for Biotechnology Information (NCBI, Bethesda, MD) and on the Internet, for use in connection with the sequence analysis programs blastp, blastn, blastx, tblastn and tblastx. By way of example, for comparisons of amino acid sequences of greater than about 30 amino acids, the Blast 2 sequences function is employed using the default BLOSUM62 matrix set to default parameters, (gap existence cost of 11, and a per residue gap cost of 1). When aligning short peptides (fewer than around 30 amino acids), the alignment is performed using the Blast 2 sequences function, employing the PAM30 matrix set to default parameters (open gap 9, extension gap 1 penalties).

An alternative indication that two nucleic acid molecules are closely related is that the two molecules hybridize to each other under stringent conditions. Stringent conditions are sequence-dependent and are different under different environmental parameters. Generally, stringent conditions are selected to be about 5° C to 20° C lower than the thermal melting point (Tm) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence remains hybridized to a perfectly matched probe or complementary strand. Conditions for nucleic acid hybridization and calculation of stringencies can be found in Sambrook *et al.* (In *Molecular Cloning: A Laboratory Manual*, CSHL, New York, 1989) and Tijssen (*Laboratory Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Acid Probes* Part I, Chapter 2, Elsevier, New York, 1993). Nucleic acid molecules that hybridize under stringent conditions to a human tyrosine kinase protein-encoding sequence will typically hybridize to a probe based on either an entire human tyrosine kinase protein-encoding sequence or selected portions of the encoding sequence under wash conditions of 2x SSC at 50° C.

Nucleic acid sequences that do not show a high degree of sequence identity may nevertheless encode similar amino acid sequences, due to the degeneracy of the genetic code. It is

understood that changes in nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid molecules that all encode substantially the same protein.

Specific binding agent: An agent that binds substantially only to a defined target. Thus a protein-specific binding agent binds substantially only the specified protein. By way of example, as used herein, the term "PDGFRA-protein specific binding agent" includes anti-PDGFRA protein antibodies (and functional fragments thereof) and other agents (such as soluble receptors) that bind substantially only to the PDGFRA protein.

Anti-PDGFRA protein antibodies (or antibodies to another tyrosine kinase) may be produced using standard procedures described in a number of texts, including Harlow and Lane (*Antibodies, A Laboratory Manual*, CSHL, New York, 1988). The determination that a particular agent binds substantially only to the specified protein may readily be made by using or adapting routine procedures. One suitable *in vitro* assay makes use of the Western blotting procedure (described in many standard texts, including Harlow and Lane (*Antibodies, A Laboratory Manual*, CSHL, New York, 1988)). Western blotting may be used to determine that a given protein binding agent, such as an anti-PDGFRA protein monoclonal antibody, binds substantially only to the PDGFRA protein.

Shorter fragments of antibodies can also serve as specific binding agents. For instance, Fabs, Fvs, and single-chain Fvs (SCFvs) that bind to a specified protein would be specific binding agents. These antibody fragments are defined as follows: (1) Fab, the fragment which contains a monovalent antigen-binding fragment of an antibody molecule produced by digestion of whole antibody with the enzyme papain to yield an intact light chain and a portion of one heavy chain; (2) Fab', the fragment of an antibody molecule obtained by treating whole antibody with pepsin, followed by reduction, to yield an intact light chain and a portion of the heavy chain; two Fab' fragments are obtained per antibody molecule; (3) (Fab')2, the fragment of the antibody obtained by treating whole antibody with the enzyme pepsin without subsequent reduction; (4) F(ab')2, a dimer of two Fab' fragments held together by two disulfide bonds; (5) Fv, a genetically engineered fragment containing the variable region of the light chain and the variable region of the heavy chain expressed as two chains; and (6) single chain antibody ("SCA"), a genetically engineered molecule containing the variable region of the light chain, the variable region of the heavy chain, linked by a suitable polypeptide linker as a genetically fused single chain molecule. Methods of making these fragments are routine.

Subject: Living multi-cellular vertebrate organisms, a category that includes both human and non-human mammals.

Target sequence: "Target sequence" is a portion of ssDNA, dsDNA or RNA that, upon hybridization to a therapeutically effective oligonucleotide or oligonucleotide analog, results in the inhibition of expression. For example, hybridization of therapeutically effectively oligonucleotide to a PDGFRA target sequence results in inhibition of PDGFRA expression. Either an antisense or a sense molecule can be used to target a portion of dsDNA, since both will interfere with the expression

of that portion of the dsDNA. The antisense molecule can bind to the plus strand, and the sense molecule can bind to the minus strand. Thus, target sequences can be ssDNA, dsDNA, and RNA.

Transformed: A transformed cell is a cell into which has been introduced a nucleic acid molecule by molecular biology techniques. As used herein, the term transformation encompasses all techniques by which a nucleic acid molecule might be introduced into such a cell, including transfection with viral vectors, transformation with plasmid vectors, and introduction of naked DNA by electroporation, lipofection, and particle gun acceleration.

Vector: A nucleic acid molecule as introduced into a host cell, thereby producing a transformed host cell. A vector may include nucleic acid sequences that permit it to replicate in a host cell, such as an origin of replication. A vector may also include one or more selectable marker genes and other genetic elements known in the art.

Unless otherwise explained, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. The singular terms "a," "an," and "the" include plural referents unless context clearly indicates otherwise. Similarly, the word "or" is intended to include "and" unless the context clearly indicates otherwise. Hence "comprising A or B" means including A, or B, or A and B. It is further to be understood that all base sizes or amino acid sizes, and all molecular weight or molecular mass values, given for nucleic acids or polypeptides are approximate, and are provided for description. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, suitable methods and materials are described below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In case of conflict, the present specification, including explanations of terms, will control. In addition, the materials, methods, and examples are illustrative only and not intended to be limiting.

III. Overview of Several Embodiments

One embodiment is an isolated variant PDGFRA polypeptide. Specific examples of such polypeptides comprise an amino acid sequence as set forth in SEQ ID NO: 4, 6, 8, 10, 12, 21, 23, or 25 or a fragment thereof comprising at least 10 contiguous amino acids including the variant site as set forth in position(s) 842 of SEQ ID NO: 4, 841 and 842 of SEQ ID NO: 6, 845 and 846 of SEQ ID NO: 8, 561 and 562 of SEQ ID NO: 10, 565 and 566 of SEQ ID NO: 12, 561 of SEQ ID NO: 21, 559 and 560 of SEQ ID NO: 23, or 841 and 842 of SEQ ID NO: 25. Also encompassed herein are the PDGFRA polypeptides defined by the consensus sequence shown in SEQ ID NO: 27, and fragments thereof, particularly fragments that overlap one or more of the noted variable regions.

Also provided are isolated nucleic acid molecules encoding such polypeptides, recombinant nucleic acid molecules comprising a promoter sequence operably linked to these nucleic acid molecules, and cells transformed with such recombinant nucleic acid molecules. Specific examples

of nucleic acid molecules comprise a nucleotide sequence as set forth in SEQ ID NO: 3, 5, 7, 9, 11, 20, 22, or 24; or a fragment thereof comprising including the variant nucleic sequence shown in position(s) 2919 of SEQ ID NO: 3, 2917 and 2918 of SEQ ID NO: 5, 2927 and 2928 of SEQ ID NO: 7, 2075 to 2080 of SEQ ID NO: 9, 2089 to 2093 of SEQ ID NO: 11, 2076 of SEQ ID NO: 20, 2017 and 2072 of SEQ ID NO: 22, or 2916 to 2919 of SEQ ID NO: 24. Also encompassed herein are the PDGFRA nucleic acid molecules defined by the consensus sequence shown in SEQ ID NO: 26, and fragments thereof, particularly fragments that overlap one or more of the noted variable regions.

A further embodiment is a method of detecting a biological condition (*e.g.*, neoplasia) associated with an activating PDGFRA mutation in a subject, comprising determining whether the subject has an activating mutation in PDGFRA, and wherein the activating mutation comprises the variant nucleic sequence shown in position(s) 2919 of SEQ ID NO: 3, 2917 and 2918 of SEQ ID NO: 5, 2927 and 2928 of SEQ ID NO: 7, 2075 to 2080 of SEQ ID NO: 9, 2089 to 2093 of SEQ ID NO: 11, 2076 of SEQ ID NO: 20, 2017 and 2072 of SEQ ID NO: 22, or 2916 to 2919 of SEQ ID NO: 24, or in any one or more of the variable positions indicated in SEQ ID NO: 26. Specific examples of biological conditions contemplated herein are neoplasias that comprise a GIST.

In specific examples of these methods, the method involves reacting at least one PDGFRA molecule contained in a clinical sample from the subject with a reagent comprising a PDGFRA-specific binding agent to form a PDGFRA:agent complex. For instance, the PDGFRA molecule in some instances is a PDGFRA encoding nucleic acid or a PDGFRA protein, and the PDGFRA specific binding agent is a PDGFRA oligonucleotide or a PDGFRA protein specific binding agent. In some embodiments, the sample from the subject includes a neoplastic cell, or is prepared from a neoplastic cell or a sample comprising a neoplastic cell.

In some of the provided methods of detecting a biological condition, the PDGFRA molecule is a PDGFRA encoding nucleic acid sequence. Specific examples of such methods involve using an agent that comprises a labeled nucleotide probe. For instance, the nucleotide probe will in some instances have a sequence as shown in SEQ ID NO: 3, 5, 7, 9, 11, 20, 22, or 24, or a fragments of one of these sequences that is at least 15 nucleotides in length, and that includes the sequence shown in position(s) 2919 of SEQ ID NO: 3, 2917 and 2918 of SEQ ID NO: 5, 2927 and 2928 of SEQ ID NO: 7, 2075 to 2080 of SEQ ID NO: 9, 2089 to 2093 of SEQ ID NO: 11, 2076 of SEQ ID NO: 20, 2017 and 2072 of SEQ ID NO: 22, or 2916 to 2919 of SEQ ID NO: 24.

Specific method embodiments involve *in vitro* amplifying a PDGFRA nucleic acid prior to detecting the activating PDGFRA mutation. By way of example, the PDGFRA nucleic acid is in some cases *in vitro* amplified using at least one oligonucleotide primer derived from a PDGFRA-protein encoding sequence, such as the specific oligonucleotide primers listed herein. Other specific oligonucleotide primers comprise at least 15 contiguous nucleotides from SEQ ID NO: 3, 5, 7, 9, 11, 20, 22, or 24. For instance, representative examples of such primers include a sequence as represented by at least 15 contiguous nucleotides shown in position(s) 2919 of SEQ ID NO: 3, 2917 and 2918 of SEQ ID NO: 5, 2927 and 2928 of SEQ ID NO: 7, 2075 to 2080 of SEQ ID NO: 9, 2089

to 2093 of SEQ ID NO: 11, 2076 of SEQ ID NO: 20, 2017 and 2072 of SEQ ID NO: 22, or 2916 to 2919 of SEQ ID NO: 24. Also included are primers that would be situated across a region including one or more of these variant positions, or any variant position indicated in SEQ ID NO: 26, so that the primers could be used to prime the amplification of a nucleic acid sequence encompassing one or more of the variants.

In other method of detection embodiments, the PDGFRA molecule is a PDGFRA protein, for instance a variant PDGFRA protein comprising a sequence as shown in SEQ ID NO: 4, 6, 8, 10, 12, 21, 23, or 25. In examples of such methods, the complexes are detected by western blot assay, or by ELISA.

Specific examples of PDGFRA-specific binding agents are PDGFRA-specific antibody or a functional fragment thereof, for instance monoclonal antibodies or fragments of monoclonal antibodies. Optionally, such monoclonal antibodies recognize an epitope of a variant PDGFRA (such as an epitope of a variant PDGFRA having an amino acid sequence as shown in SEQ ID NO: 4, 6, 8, 10, 12, 21, 23, or 25) and not (or to a lesser extent) an epitope of wildtype PDGFRA. In particular methods, the antibody is reactive to an epitope including the amino acid sequence shown in position(s) 842 of SEQ ID NO: 4, 841 and 842 of SEQ ID NO: 6, 845 and 846 of SEQ ID NO: 8, 561 and 562 of SEQ ID NO: 10, 565 and 566 of SEQ ID NO: 12, 561 of SEQ ID NO: 21, 559 and 560 of SEQ ID NO: 23, or 841 and 842 of SEQ ID NO: 25.

Also provided in the disclosure are kits for detecting an activating PDGFRA mutation in a subject using methods described herein. Examples of such kits are used with protein-detection methods, and include at least one PDGFRA protein specific binding agent. For instance, in specific kits the agent (*e.g.*, an antibody) is capable of specifically binding to an epitope within a PDGFRA variant protein but not to an epitope of wildtype PDGFRA. Thus, some such agents are capable of specifically binding to an epitope within the amino acid sequence shown in SEQ ID NO: 4, 6, 8, 10, 12, 21, 23, or 25, or more particularly antigenic fragments of (a) that comprise the sequence shown in position(s) 842 of SEQ ID NO: 4, 841 and 842 of SEQ ID NO: 6, 845 and 846 of SEQ ID NO: 8, 561 and 562 of SEQ ID NO: 10, 565 and 566 of SEQ ID NO: 12, 561 of SEQ ID NO: 21, 559 and 560 of SEQ ID NO: 23, or 841 and 842 of SEQ ID NO: 25. Examples of the protein-detection kits further include a means for detecting binding of the PDGFRA protein binding agent to a PDGFRA polypeptide.

A further embodiment is a kit for determining whether or not a subject (*e.g.*, an animal, or more particularly a mammal) has a biological condition (*e.g.*, neoplasia, such as that comprising a GIST) associated with an activating PDGFRA mutation by detecting a mutant PDGFRA sequence in the subject, which kit includes a container comprising at least one oligonucleotide specific for a PDGFRA mutation sequence; and instructions for using the kit, the instructions indicating steps for performing a method to detect the presence of mutant PDGFRA nucleic acid in the sample; and analyzing data generated by the method, wherein the instructions indicate that presence of the mutant nucleic acid in the sample indicates that the individual has or is predisposed to the biological

condition. Optionally, such kits further include at least one container that comprises a detectable oligonucleotide. Specific examples of oligonucleotides (labeled or not) that may be included in these kits will be specific for a PDGFRA mutation sequence. For instance, particular example oligonucleotides comprise a sequence specific for a PDGFRA encoding sequence and containing the specific sequence shown in shown in position(s) 2919 of SEQ ID NO: 3, 2917 and 2918 of SEQ ID NO: 5, 2927 and 2928 of SEQ ID NO: 7, 2075 to 2080 of SEQ ID NO: 9, 2089 to 2093 of SEQ ID NO: 11, 2076 of SEQ ID NO: 20, 2017 and 2072 of SEQ ID NO: 22, or 2916 to 2919 of SEQ ID NO: 24.

Another specific embodiment is a kit for determining whether or not a subject (*e.g.*, an animal, or more particularly a mammal) has a biological condition (*e.g.*, neoplasia, such as that comprising a GIST) associated with an activating PDGFRA mutation, the kit including a container comprising a PDGFRA mutant specific antibody; a container comprising a negative control sample; and instructions for using the kit, the instructions indicating steps for: performing a test assay to detect a quantity of PDGFRA mutant protein in a test sample of tissue and/or bodily fluid from the subject, performing a negative control assay to detect a quantity of PDGFRA mutant protein in the negative control sample; and comparing data generated by the test assay and negative control assay, wherein the instructions indicate that a quantity of PDGFRA mutant protein in the test sample more than the quantity of PDGFRA mutant protein in the negative control sample indicates that the subject has the biological condition. Specific examples of such kits further include one or more detectable antibodies that bind to the antibody specific for PDGFRA mutant protein (e.g., to be used in detection of the primary antibody).

Yet another embodiment is a method of screening for a compound useful in influencing (for instance, inhibiting or treating) PDGFRA-mediated neoplasia in a mammal, comprising determining if a test compound binds to or interacts with the polypeptide or fragment according to claim 1, and selecting a compound that so binds. In specific examples of this method, binding of the compound inhibits a PDGFRA protein biological activity (*e.g.*, kinase activity). In certain examples, the test compound is applied to a test cell. Compounds identified or selected by such methods, whether or not formulated for use as therapeutic agents, are also contemplated.

Also provided are compositions that include at least one antigenic fragment of a provided PDGFRA variant protein, where the antigenic fragment includes the variant sequence as shown at position(s) 842 of SEQ ID NO: 4, 841 and 842 of SEQ ID NO: 6, 845 and 846 of SEQ ID NO: 8, 561 and 562 of SEQ ID NO: 10, 565 and 566 of SEQ ID NO: 12, 561 of SEQ ID NO: 21, 559 and 560 of SEQ ID NO: 23, or 841 and 842 of SEQ ID NO: 25.

IV. Identification of Activating Mutations of PDGFRA

The inventors have determined that mutations in the platelet derived growth factor receptor alpha (PDGFRA) gene, particularly mutations that produce activated PDGFRA protein, are linked to neoplastic disease such as cancer, and thereby can be used to assess whether a subject suffers from or

is susceptible to such a condition. The following examples illustrate this by showing particular examples of mutations that are associated with specific cancers in human subjects. Moreover, guidance is provided about finding other mutations associated with other specific cancers, both in PDGFRA and in other tyrosine kinases. Hence, in its broadest aspect, the disclosure is not limited to particular mutations, but is instead premised on the finding that activating PDGFRA mutations are associated with neoplastic disease.

The PDGFRA protein is a type III receptor tyrosine kinase with homology to KIT, FLT3, CSF1-R, and PDGFR beta (PDGFRB). Although PDGFRA activation has been hypothesized to be involved in certain cancers, most notably gliomas, evidence of genomic activation in human cancer has only recently been reported in two cases of myeloproliferative disease associated with translocation of the BCR and PDGFRA genes. We report herein several novel mutations of PDGFRA resulting in constitutive activation. These mutations were initially discovered in GISTs. Based on experience with KIT and FLT3, it is likely that mutations in other regions of the PDGFRA gene may result in constitutive activation of tyrosine kinase activity. At least in the case of KIT, the site of mutation varies between different diseases (*e.g.*, mastocytosis vs. GIST). Finally, findings reported herein strongly suggest that similar mutations can activate related family members PDGFRB and CSF-1R, and that these mutant proteins are likely to be therapeutic targets in human cancer.

The discovery that mutations in the sequence of *PDGFRA* predisposes a subject to developing neoplasms also enables a variety of diagnostic, prognostic, and therapeutic methods that are further embodiments. The new appreciation of the role of activated PDGFRA in neoplastic diseases, such as cancers, enables detection of predisposition to or diagnosis of these conditions in a subject. This disclosure also enables early detection of subjects at high risk of these conditions, identification of subjects with particularly severe disease and/or tendency to progress, and in some embodiments detection of resistance or susceptibility of a subject to drug(s). Identification of the activating mutations described herein provides opportunities for prevention and/or early treatment as well as particular treatment selection.

V. Diagnostic and Therapeutic Applications

The presence of *PDGFRA* gene mutations in GIST strongly suggests that other human cancers will have similar mutations. When present in a cancer, mutant isoforms of PDGFRA represent a therapeutic target for tyrosine kinase inhibitors (TKIs), immunotherapy and other novel targeted approaches. Because *PDGFRA* gene mutations are not found in all tumors, the selection of patients for therapy targeting mutant PDGFRA isoforms would be optimized by pre-therapy analysis of cancer cells for the presence of *PDGFRA* gene mutations.

Such analysis can be based on PCR-based assays for these mutations, using for instance one or more of the following approaches: size fractionation by gel electrophoresis, direct sequencing, single-strand conformation polymorphism (SSCP), high pressure liquid chromatography (including partially denaturing HPLC), allele-specific hybridization, amplification refractory mutation

screening, *PDGFRA* mutation screening by oligonucleotide microarray, restriction fragment polymorphism, MALDI-TOF mass spectrometry, or various related technologies (Abu-Duhier *et al.*, *Br. J Haematol.*, 113: 983-988, 2001; Kottaridis *et al.*, *Blood*, 98: 1752-1759, 2001; Choy *et al.*, *Ann. Hum. Gen.*, 63: 383-391, 1999; Grompe, *Nature Genetics*, 5: 111-117, 1993; Perlin & Szabady, *Hum. Mutat.*, 19: 361-373, 2002; Amos & and Patnaik, *Hum. Mutat.*, 19: 324-333, 2002; Cotton, *Hum. Mutat.*, 19: 313-314, 2002; Stirewalt *et al.*, *Blood*, 97: 3589-3595, 2001; Hung *et al.*, *Blood Coagul. Fibrinolysis*, 13: 117-122, 2002; Larsen *et al.*, *Pharmacogenomics*, 2: 387-399, 2001; Shchepinov *et al.*, *Nucleic Acids Res.*, 29: 3864-3872, 2001).

In addition, mutant PDGFRA proteins may be detected through novel epitopes recognized by polyclonal and/or monoclonal antibodies used in ELISA, immunoblotting, flow cytometric, immunohistochemical and other mutant protein detection strategies (Wong *et al.*, *Cancer Res.*, 46: 6029-6033, 1986; Luwor *et al.*, *Cancer Res.*, 61: 5355-5361, 2001; Mishima *et al.*, *Cancer Res.*, 61: 5349-5354, 2001; Ijaz *et al.*, *J. Med. Virol.*, 63: 210-216, 2001). Additionally mutant PDGFRA proteins could be detected by mass spectrometry assays coupled to immunaffinity assays, the use of matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass mapping and liquid chromatography/quadrupole time-of-flight electrospray ionization tandem mass spectrometry (LC/Q-TOF-ESI-MS/MS) sequence tag of tumor derived proteins separated by two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) (Kiernan *et al.*, *Anal. Biochem.*, 301: 49-56, 2002; Poutanen *et al.*, *Mass Spectrom.*, 15: 1685-1692, 2001). All of these approaches may be used to detect a sequence anomaly or variant of the PDGFRA protein, a relative increase in the phosphorylation of the protein, or an increase in the inherent kinase activity of the protein.

In addition to direct detection of mutant PDGFRA proteins, it is expected that various PDGFRA mutants will result in distinctive signal transduction profiles that could be detected by global gene expression profile or analysis of the activation of various signaling intermediates (*e.g.*, STAT5) (Hofmann *et al.*, *Lancet*, 359: 481-486, 2002).

Utility of this disclosure is highlighted by the correlative studies of response to imatinib mesylate and tumor KIT genotype in patients treated in a phase II trial of imatinib mesylate. In this trial, response to treatment was vastly superior in patients with an imatinib mesylate-sensitive KIT mutation compared with patients with no detectable KIT mutation (Heinrich *et al.*, *Proc. of ASCO*, 21:2A, 2002).

It is believed that the nature and location of PDGFRA mutations affects the sensitivity of the resultant mutant protein to various TKIs. For example, imatinib mesylate is highly active against the kinase activity of wild-type KIT and against activating mutations involving the extracellular, juxtamembrane and TK1 domain (Tuveson *et al.*, *Oncogene*, 20: 5054-5058, 2001; Heinrich *et al.*, *Blood*, 96: 925-932, 2000). In contrast, imatinib mesylate has no clinically useful activity against mutations of the aspartic acid residue at position 816 (*e.g.*, D816V, D816Y, D816F, or D816H) (Ma *et al.*, *Blood*, 99: 1741-1744, 2002). The KIT D816V mutation is homologous to the D842V PDGFRA mutation described in this application. In addition, indolinone and tyrphostin compounds

have little or no activity against KIT D816 mutations (or the equivalent D814 residue in murine KIT) but are potent inhibitors of the kinase activity of wild-type and juxtamembrane mutant KIT polypeptides (Ma *et al.*, *Blood*, 99: 1741-1744, 2002; Ikeda *et al.*, *Blood* 96, 99a-99a. 11-16-2000; Ma *et al.*, *J. Invest. Derma.*, 114: 392-394, 2000). However, imatinib mesylate has some activity against other KIT activation loop mutations that involve residues other than aspartic acid 816.

Based on homology to KIT, it is predicted that imatinib mesylate and indolinone compounds would have minimal activity against the D842V PDGFRA mutation but might have clinically useful activity against PDGFRA deletion and/or insertion mutations. In the absence of structural biology information concerning the structure of both wild type and mutant PDGFRA proteins and the site of binding of imatinib mesylate or other TKIs to these proteins, it will be necessary to empirically determine the activity of TKIs against the kinase activity of various mutant PDGFRA proteins. This could be accomplished by cloning cDNAs of the various PDGFRA mutant isoforms and the recombinant protein in prokaryotic or eukaryotic cells (Ma *et al.*, *Blood*, 99: 1741-1744, 2002; Wood *et al.*, *Cancer Res*, 60: 2178-2189, 2000). Protein expressed in such a manner could be used to determine biochemical activity of existing TKIs and could also be used in high throughput screening of chemical libraries to help identify and optimize pre-clinical development of new compounds against these or other PDGFRA mutant isoforms (Chroeder *et al.*, *J. Med. Chem.*, 44: 1915-1926, 2001; Hamby *et al.*, *J. Med. Chem.*, 40: 2296-2303, 1997; Druker *et al.*, *Nature Medicine,* 2: 561-566, 1996). Prior determination of biochemical potency of specific compounds to different PDGFRA mutations would allow clinical testing of patient specimens for PDGFRA mutations and selection of the appropriate TKI based on the specific mutation and sensitivity associated with that patient's tumor.

Since the novel PDGFRA activating protein variants are only expressed by neoplastic cells, they have the potential to serve as tumor-specific antigens for cytotoxic T-lymphocytes (CTL). Indeed, it has been shown that the unique peptide sequence generated by the BCR-ABL fusion protein characteristic of chronic myelogenous leukemia can serve as the basis of an in vivo immune therapy that utilizes BCR-ABL peptide loaded dendritic cells to generate CTL with BCR-ABL specificity (He *et al.*, *Cancer Immunol. Immunother.*, 50: 31-40, 2001).

VI. Prediction of Additional Types of PDGFRA Mutations

Based on experience with KIT and FLT3, it is likely that mutations in other regions of the PDGFRA gene may result in constitutive activation of tyrosine kinase activity. Other likely sites of PDGFRA activating mutations include the proximal extra-cellular, juxtamembrane, and TK1 domains of PDGFRA (Rubin *et al.*, *Cancer Res*, 61: 8118-8121, 2001; Lux *et al.*, *Am.J.Pathol.*, 156: 791-795, 2000; Abu-Duhier *et al.*, *Br. J. Haematol.*, 111: 190-195, 2000). Indeed, it should be noted that there is one solitary case report of an astrocytoma with a large in-frame deletion of 81 amino acids involving portions of the fourth and fifth immunoglobulin domains of PDGFRA. The tumor in that report had genomic amplification of this PDGFRA mutant allele. The activity of PDGFRA kinase of

this mutant isoform was not reported (Kumabe *et al.*, *Oncogene*, 7: 627-633, 1992). Recently Baxter *et al.* reported a translocation having the structure t(4;22)(q12;q11) in two cases of atypical chronic myeloid leukemia. Molecular cloning of the translocation revealed fusion of a portion of the BCR gene with part of exon 12 of PDGFRA (Baxter *et al.*, *Hum. Mol. Genet.* 11:1391-1397, 2002). The fusion gene from these translocations is predicted to encode a constitutively activated tyrosine kinase, however no formal biochemical characterization of these proteins was performed (Baxter *et al.*, 2002). Without meaning to be limited to a single interpretation, it is believed that fusion mechanisms of oncogenesis involving PDGFRA (*e.g.*, the BCR-PDGFRA fusions reported by Baxter *et al.*) likely are a rare occurrence, while point mutation and deletion activations are expected to be more common, and that these two mechanisms are independent of each other.

In KIT, FLT3, and CSF-1R, kinase activation results from a variety of amino acid substitutions at the conserved aspartic acid in the activation loop (D816 KIT, D835 FLT3, and D802 of CSF-1R) (Morley *et al.*, *Oncogene*, 18: 3076-3084, 1999; Moriyama *et al.*, *J. Biol. Chem.*, 271: 3347-3350, 1996). In the case of KIT and FLT3, a number of these substitutions have been found in association with certain malignancies (Ma *et al.*, *Blood*, 99: 1741-1744, 2002; Abu-Duhier *et al.*, *Br. J Haematol.*, 113: 983-988, 2001; Yamamoto *et al.*, *Blood*, 97: 2434-2439, 2001; Longley *et al.*, *Leuk. Res.*, 25: 571-576, 2001; Ning *et al.*, *Leuk. Lymphoma*, 41: 513-522, 2001). To date, no mutations of D802 of CSF-1R have been found in any human cancer. Thus far, we have found only a valine substitution at D842 of PDGFRA, but it can be predicted that a variety of amino acid substitutions at this position of PDGFRA would be activating. Assuming a single nucleotide change in codon 842, the most likely possible mutations of PDGFRA would be substitution of Asparagine, Tyrosine, Histidine, Valine, Alanine, Glycine, or Glutamic acid for the normal Aspartic acid. We predict that these additional PDGFRA mutations would also be oncogenic and will be found in one or more human neoplasms.

VII. Prediction of Similar Activating Mutations in PDGFRB

The amino acid sequence of the members of the Type III receptor tyrosine kinase family are highly conserved in the activation loop:

DFGLAR**D**IMHDSN	Human PDGFRA
DFGLAR**D**IMRDSN	Human PDGFRB
DFGLAR**D**IKNDSN	Human KIT
DFGLAR**D**IMNDSN	Human CSF-1R
DFGLAR**D**IMSDSN	Human FLT3

As noted above, amino acid substitutions at the conserved aspartic acid (shown in bold) result in constitutive activation of the tyrosine kinase activity of KIT, PDGFRA or FLT3 in different human malignancies (Rosnet *et al.*, *Blood*, 82: 1110-1119, 1993; Claesson-Welsh *et al.*, *Proc. Natl. Acad.*

Sci. U.S.A, 86: 4917-4921, 1989; Gronwald *et al.*, *Proc. Natl. Acad. Sci. U.S.A*, 85: 3435-3439, 1988; Yarden *et al.*, *Nature*, 323: 226-232, 1986). Amino acid substitution at the same aspartic acid of CSF-1R is also activating, but has not yet been found in association with human disease. Based on our findings, we predict that amino acid substitution at the same aspartic acid of PDGFRB would also be activating and that this mutation will be found in some human malignances.

VIII. Identification of Compounds that Inhibit PDGFRA Variants

This disclosure further relates in some embodiments to novel methods for screening test compounds for their ability to treat, detect, analyze, ameliorate, reverse, and/or prevent neoplasia, especially pre-cancerous lesions. In particular, the present disclosure provides methods for identifying test compounds that can be used to treat, ameliorate, reverse, and/or prevent neoplasia, including precancerous lesions. The compounds of interest can be tested by exposing the novel activating PDGFRA variants described herein to the compounds, and if a compound inhibits one of the PDGFRA variants, the compound is then further evaluated for its anti-neoplastic properties.

One aspect involves a screening method to identify a compound effective for treating, preventing, or ameliorating neoplasia, which method includes ascertaining the compound's inhibition of a provided novel activating PDGFRA variant or another activating PDGFRA variant. In some embodiments, the screening method further includes determining whether the compound inhibits the growth of tumor cells in a cell culture.

By screening compounds in this fashion, potentially beneficial and improved compounds for treating neoplasia can be identified more rapidly and with greater precision than possible in the past.

A. In General

Activating tyrosine kinase mutants, for instance the novel activating PDGFRA variants described herein, are useful to identify compounds that can be used to treat, ameliorate, or prevent neoplasms.

The screening or creation, identification and selection of appropriate high affinity inhibitors of activating PDGFRA mutants can be accomplished by a variety of methods. Broadly speaking these may include, but are not limited to, two general approaches. One approach is to use structural knowledge about the target enzyme to design a candidate molecule with which it will precisely interact. An example would be computer assisted molecular design. A second approach is to use combinatorial or other libraries of molecules, whereby a large library of molecules is screened for affinity with regard to the target enzyme.

Cancer and precancer may be thought of as diseases that involve unregulated cell growth. Cell growth involves a number of different factors. One factor is how rapidly cells proliferate, and another involves how rapidly cells die. Cells can die either by necrosis or apoptosis depending on the type of environmental stimuli. Cell differentiation is yet another factor that influences tumor growth kinetics. Resolving which of the many aspects of cell growth a test compound affects can be

important to the discovery of a relevant target for pharmaceutical therapy. Screening assays based on this technology can be combined with other tests to determine which compounds have growth inhibiting and pro-apoptotic activity.

B. Inhibitor Screening

Some embodiments provided herein involve determining the ability of a given compound to inhibit activating PDGFRA mutants, for instance the ability to specifically inhibit constitutive kinase and/or transforming activities in the PDGFRA D842V, PDGFRA V561D, PDGFRA DIMH842-845, PDGFRA HSDN845-848P, insertion ER561-562, or SPDGHE566-571R, RD841-842KI, or RVIES560-564 deletion mutants described herein. Test compounds can be assessed for their probable ability to treat neoplastic lesions either directly, or indirectly by comparing their activities against compounds known to be useful for treating neoplasia. In particular, the compounds are tested for their ability to inhibit a neoplasia that is found to contain an activating PDGFRA mutation.

C. Determining Tyrosine Kinase Influencing Activity

Compounds can be screened for inhibitory or other effects on the activity of the novel activating PDGFRA mutants described herein using an expressed recombinant version of the enzyme, or a homolog or ortholog isolated from another species. Alternatively, cells expressing one of these tyrosine kinases can be treated with a test compound and the effect of the test compound on phosphorylation of a specific target can be determined, for instance using one of the techniques described herein. Additional detail regarding methods for determining tyrosine kinase phosphorylation influencing activity (*e.g.*, inhibition) is provided herein.

D. Determining Whether a Compound Reduces the Number of Tumor Cells

In an alternate embodiment, provided screening methods involve further determining whether the compound reduces the growth of tumor cells, for instance tumor cells known to express an activated tyrosine kinase mutation such as a mutation in PDGFRA.

Various cell lines can be used, which may be selected based on the tissue to be tested. For example, these cell lines include: SW-480 -- colonic adenocarcinoma; HT-29 -- colonic adenocarcinoma, A-427 -- lung adenocarcinoma carcinoma; MCF-7--breast adenocarcinoma; and UACC-375 -- melanoma line; and DU145 -- prostate carcinoma. Cell lines can also be used that are known to express activated, mutant, tyrosine kinase proteins, for example: GIST882 – gastrointestinal stromal tumor cell line expressing KIT tyrosine kinase point mutant; SKBR3 – breast carcinoma cell line expressing ERBB2 amplification mutant; and K562 – leukemia cell line expressing BCR-ABL tyrosine kinase fusion mutant. Cytotoxicity data obtained using these cell lines are indicative of an inhibitory effect on neoplastic lesions. Certain cell lines are well characterized, and are used for instance by the United States National Cancer Institute (NCI) in their screening program for new anti-cancer drugs. Though a compound may be identified by its ability to inhibit a specific tyrosine kinase activating mutant, its activity likely will not be limited to inhibition of only that mutant protein, thus testing in different cell lines and samples is beneficial to determine the scope of its activity.

By way of example, a test compound's ability to inhibit tumor cell growth *in vitro* can be measured using the HT-29 human colon carcinoma cell line obtained from ATCC (Bethesda, MD). HT-29 cells have previously been characterized as a relevant colon tumor cell culture model (Fogh & Trempe, In: *Human Tumor Cells in Vitro*, Fogh (ed.), Plenum Press, N.Y., pp. 115-159, 1975). HT-29 cells are maintained in RPMI media supplemented with 5% fetal bovine calf serum (Gemini Bioproducts, Inc., Carlsbad, Calif.) and 2 mM glutamine, and 1% antibiotic-antimycotic, in a humidified atmosphere of 95% air and 5% CO_2 at 37° C. Briefly, HT-29 cells are plated at a density of 500 cells/well in 96 well microtiter plates and incubated for 24 hours at 37° C. prior to the addition of test compound. Each determination of cell number involved six replicates. After six days in culture, the cells are fixed by the addition of cold trichloroacetic acid (TCA) to a final concentration of 10% and protein levels are measured, for instance using the sulforhodamine B (SRB) colorimetric protein stain assay as previously described by Skehan *et al.* (*J. Natl. Cancer Inst.* 82: 1107-112, 1990). In addition to the SRB assay, a number of other methods are available to measure growth inhibition and could be substituted for the SRB assay. These methods include counting viable cells following trypan blue staining, labeling cells capable of DNA synthesis with BrdU or radiolabeled thymidine, neutral red staining of viable cells, or MTT staining of viable cells.

Significant tumor cell growth inhibition greater than about 30% at a dose of 100 μM or below is further indicative that the compound is useful for treating neoplastic lesions. An IC_{50} value may be determined and used for comparative purposes. This value is the concentration of drug needed to inhibit tumor cell growth by 50% relative to the control. In some embodiments, the IC_{50} value is less than 100 μM in order for the compound to be considered further for potential use for treating, ameliorating, or preventing neoplastic lesions.

E. Determining Whether a Test Compound Induces Apoptosis

In other embodiments, screening methods provided herein further involve determining whether the test compound induces apoptosis in cultures of tumor cells.

Two distinct forms of cell death may be described by morphological and biochemical criteria: necrosis and apoptosis. Necrosis is accompanied by increased permeability of the plasma membrane, whereby the cells swell and the plasma membrane ruptures within minutes. Apoptosis is characterized by membrane blebbing, condensation of cytoplasm, and the activation of endogenous endonucleases.

Apoptosis occurs naturally during normal tissue turnover and during embryonic development of organs and limbs. Apoptosis also can be induced by various stimuli, including cytotoxic T-lymphocytes and natural killer cells, by ionizing radiation and by certain chemotherapeutic drugs. Inappropriate regulation of apoptosis is thought to play an important role in many pathological conditions including cancer, AIDS, or Alzheimer's disease, etc.

Test compounds can be screened for induction of apoptosis using cultures of tumor cells maintained under conditions as described above. In some examples of such screening methods, treatment of cells with test compounds involves either pre- or post-confluent cultures and treatment

for two to seven days at various concentrations of the test compounds. Apoptotic cells can be measured in both the attached and "floating" portions of the cultures. Both are collected by removing the supernatant, trypsinizing the attached cells, and combining both preparations following a centrifugation wash step (10 minutes, 2000 rpm). The protocol for treating tumor cell cultures with sulindac and related compounds to obtain a significant amount of apoptosis has been described in the literature (*e.g.*, Piazza *et al.*, *Cancer Res.*, 55:3110-16, 1995). Particular features include collecting both floating and attached cells, identification of the optimal treatment times and dose range for observing apoptosis, and identification of optimal cell culture conditions.

Following treatment with a test compound, cultures can be assayed for apoptosis and necrosis, for instance by florescent microscopy following labeling with acridine orange and ethidium bromide. Many methods for measuring apoptotic cells are known to those of ordinary skill in the art; for instance, one method for measuring apoptotic cell number has been described by Duke & Cohen (*Curr. Prot. Immuno.*, Coligan *et al.*, eds., 3.17.1-3.17.1, 1992).

For example, floating and attached cells are collected by trypsinization and washed three times in PBS. Aliquots of cells are then centrifuged. The pellet is resuspended in media and a dye mixture containing acridine orange and ethidium bromide prepared in PBS and mixed gently. The mixture then can be placed on a microscope slide and examined for morphological features of apoptosis.

Apoptosis also can be quantified by measuring an increase in DNA fragmentation in cells that have been treated with test compounds. Commercial photometric EIA for the quantitative *in vitro* determination of cytoplasmic histone-associated-DNA-fragments (mono- and oligo-nucleosomes) are available (*e.g.*, Cell Death Detection ELISA, Boehringer Mannheim). The Boehringer Mannheim assay is based on a sandwich-enzyme-immunoassay principle, using mouse monoclonal antibodies directed against DNA and histones, respectively. This allows the specific determination of mono- and oligo-nucleosomes in the cytoplasmic fraction of cell lysates. According to the vendor, apoptosis is measured as follows: The sample (cell-lysate) is placed into a streptavidin-coated microtiter plate ("MTP"). Subsequently, a mixture of anti-histone-biotin and anti-DNA peroxidase conjugates is added and incubated for two hours. During the incubation period, the anti-histone antibody binds to the histone-component of the nucleosomes and simultaneously fixes the immunocomplex to the streptavidin-coated MTP via its biotinylation. Additionally, the anti-DNA peroxidase antibody reacts with the DNA component of the nucleosomes. After removal of unbound antibodies by a washing step, the amount of nucleosomes is quantified by the peroxidase retained in the immunocomplex. Peroxidase is determined photometrically with ABTS7 (2,2'-Azido-[3-ethylbenzthiazolin-sulfonate]) as substrate.

By way of example, SW-480 colon adenocarcinoma cells are plated in a 96-well MTP at a density of 10,000 cells per well. Cells are then treated with test compound, and allowed to incubate for 48 hours at 37° C. After the incubation, the MTP is centrifuged and the supernatant is removed. The cell pellet in each well is then resuspended in lysis buffer for 30 minutes. The lysates are then

centrifuged and aliquots of the supernatant (*i.e.*, cytoplasmic fraction) are transferred into a streptavidin-coated MTP. Care is taken not to shake the lysed pellets (*i.e.*, cell nuclei containing high molecular weight, un-fragmented DNA) in the MTP. Samples are then analyzed. Fold stimulation ($FS = OD_{max} / OD_{veh}$), an indicator of apoptotic response, is determined for each compound tested at a given concentration. EC_{50} values may also be determined by evaluating a series of concentrations of the test compound.

Statistically significant increases of apoptosis (*i.e.*, greater than 2 fold stimulation at a test compound concentration of 100 μM) are further indicative that the compound is useful for treating neoplastic lesions. Preferably, the EC_{50} value for apoptotic activity should be less than 100 μM for the compound to be further considered for potential use for treating neoplastic lesions. EC_{50} is understood herein to be the concentration that causes 50% induction of apoptosis relative to vehicle treatment.

F. Organ Culture Model Tests

Test compounds identified by the methods described herein can be tested for antineoplastic activity by their ability to inhibit the incidence of preneoplastic lesions in an organ culture system, such as a mammary gland organ culture system. The mouse mammary gland organ culture technique has been successfully used by other investigators to study the effects of known antineoplastic agents such as NSAIDs, retinoids, tamoxifen, selenium, and certain natural products, and is useful for validation of the screening methods provided herein.

By way of example, female BALB/c mice can be treated with a combination of estradiol and progesterone daily, in order to prime the glands to be responsive to hormones *in vitro*. The animals are sacrificed, and thoracic mammary glands are excised aseptically and incubated for ten days in growth media supplemented with insulin, prolactin, hydrocortisone, and aldosterone. DMBA (7,12-dimethylbenz(a)anthracene) is added to medium to induce the formation of premalignant lesions. Fully developed glands are then deprived of prolactin, hydrocortisone, and aldosterone, resulting in the regression of the glands but not the premalignant lesions.

The test compound is dissolved in, for instance, DMSO and added to the culture media for the duration of the culture period. At the end of the culture period, the glands are fixed in 10% formalin, stained with alum carmine, and mounted on glass slides. The incidence of forming mammary lesions is the ratio of the glands with mammary lesions to glands without lesions. The incidence of mammary lesions in test compound treated glands is compared with that of the untreated glands.

The extent of the area occupied by the mammary lesions can be quantitated by projecting an image of the gland onto a digitation pad. The area covered by the gland is traced on the pad and considered as 100% of the area. The space covered by each of the unregressed structures is also outlined on the digitization pad and quantitated by the computer.

The following examples are provided to illustrate certain particular features and/or embodiments. These examples should not be construed to limit the invention to the particular features or embodiments described.

EXAMPLES

The PDGFRA protein is a type III receptor tyrosine kinase with homology to KIT, FLT3, CSF1-R and PDGFR beta (PDGFRB). Although PDGFRA activation has been suspected to be involved in certain cancers, most notably gliomas, evidence of genomic activation in human cancer has not been previously reported. Provided herein are novel mutations of PDGFRA resulting in constitutive activation. These mutations were initially discovered in GISTs. It is expected that other human cancers will have identical or similar mutations. Based on experience with KIT and FLT3, it is likely that mutations in other regions of the PDGFRA gene may result in constitutive activation of tyrosine kinase activity. At least in the case of KIT, the site of mutation varies between different diseases (*e.g.*, mastocytosis vs. GIST). Finally, these findings strongly suggest that similar mutations can activate related family members PDGFRB and CSF-1R, and that these mutant proteins are likely to be therapeutic targets in human cancer.

EXAMPLE 1: Activating mutations in PDGFRA in GISTs

Methods

Three to five mm^3 pieces of frozen gastrointestinal stromal tumors were homogenized by 5 to 10 strokes of a Tissue Tearor™ homogenizer in ice-cold lysis buffer (1% Nonidet P-40, 50 mmol/L Tris, pH 8.0, 100 mmol/L sodium fluoride, 30 mmol/L sodium pyrophosphate, 2 mmol/L sodium molybdate, 5 mmol/L ethylenediaminetetracetic acid, 2 mmol/L sodium vanadate, 10μg/ml aprotinin, 10μg/ml leupeptin, and 100 μg/ml phenylmethylsulfonyl fluoride) and rocked overnight at 4 °C. Residual cell debris was removed by centrifugation (14,000g) for 20 minutes at 4 °C, and the supernatant protein concentrations were determined using the BioRad™ MMT assay. Five hundred microliters (μl) of protein cell lysates (2 mg/ml) were pre-cleared with 20μl of normal rabbit serum (Zymed Laboratories) and 20μl of protein A sepharose 4B (Zymed Laboratories) for one hour at 4 °C, followed by sequential additions of 20 μl of panRTK antibodies and 20μl of protein A sepharose 4B with end-to-end rotation for two hours after each addition.

Antibodies used for immunoprecipitation were to KIT (Santa Cruz sc-168), PDGFRA (Santa Cruz sc-338), and panRTK. The panRTK antibodies were raised against combinations of epitopes, each epitope representing one variation of the conserved RTK catalytic domain sequence (#1 YVHRDLAARNIL (SEQ ID NO: 13); #2 CIHRDLAARNVL (SEQ ID NO: 14); #3 FVHRDLAARNCM (SEQ ID NO: 15); #4 LVHRDLAARNVL (SEQ ID NO: 16); #5 FIHRDIAARNCL (SEQ ID NO: 17); and #6 FVHRDLATRNCL (SEQ ID NO: 18)). Each rabbit was injected with three panRTK epitopes, either combination #1 (YVHRDLAARNIL, CIHRDLAARNVL and FVHRDLAARNCM) or combination #2 (LVHRDLAARNVL,

FIHRDIAARNCL, and FVHRDLATRNCL). The panRTK antisera were then affinity purified using the same combinations of epitopes against which they had been raised. These panRTK antisera are expected to react with all human and murine RTKs, and with a subset of nonreceptor tyrosine kinase proteins (*e.g.*, JAK family members, SRC family members, FAK/PTK2, ABL, and ARG) that contain the conserved epitope. The panRTK antisera immunoprecipitate individual RTK proteins with lower efficiency than specific kinase antibodies, inasmuch as they react with the entire class of RTK proteins, rather than targeting a specific kinase protein. Typically, 10 – 20 μg of panRTK antisera are required per immunoprecipitation, in order to purify the same amount of each RTK protein that would typically be immunoprecipitated with 2 – 4 μg of an optimized, specific antibody.

The immunoprecipitates were then washed three times in lysis buffer, 10 minutes each wash, and once in 10 mM Tris for one hour. After discharging the supernatant, 20μl of sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis loading buffer was added to the immunoprecipitates, and heated for six minutes at 95 °C. The supernatants were then collected and loaded into 4-12% sodium dodecyl sulfate-polyacrylamide gel gradient gels (NuPAGE™, Invitrogen, Carlsbad, CA), followed by electrophoretic transfer to nitrocellulose membranes (PROTRAN™, Schleicher & Schuell, Keene, NH). Ponceau S solution was used to confirm adequate protein transfer (Sigma Chemical Co., St Louis, MO).

The membranes were then blocked overnight using a 1% solution of bovine serum albumin (BSA; Sigma Chemical Co., St Louis, MO) in 0.01% phosphate-buffered saline (PBS)-Tris at pH 7.4. Protein tyrosine phosphorylation was detected by staining the membranes with anti-phosphotyrosine monoclonal mouse antibody (PY99; Santa Cruz Biotechnology, Santa Cruz, CA; 1: 4000) in 1% BSA/0.01 % PBS-Tris solution for 2 hours at room temperature (RT) and with anti-mouse immunoglobulin-horseradish peroxidase goat polyclonal antibody (Amersham Pharmacia Biotech, Piscataway, NJ; 1:5000). The membranes were then stripped, blocked with 5% non-fat milk/0.01% PBS-Tris solution for one hour at room temperature, and restained with specific antibodies to PDGFRA (Santa Cruz) or KIT (Dako). All antibody reactions were detected by chemiluminescence (ECL; Pierce, Rockford, IL).

Tumor tissue was identified on unstained, 5 μm sections by comparison with H&E (Hematoxylin and Eosin) stained slides and was carefully collected using a clean, sterile scalpel blade into a microfuge tube. Dissection by this approach was straightforward and there was minimal contamination from adjacent normal tissue. Dissected tissue was deparaffinized by serial extraction with xylene and ethanol and allowed to air-dry. DNA was extracted using the Qiagen mini-kit (Qiagen, 51304) in accordance with the manufacturer's recommendations.

0.5 μg of purified tumor DNA was subjected to 45 cycles of *in vitro* amplification by polymerase chain reaction (PCR) using the High Fidelity PCR System (Roche #1732078). Primer pairs for each exon analyzed are listed in Table 1. Negative controls were included in every set of amplifications. In a minority of cases there was insufficient amplified DNA for screening by HPLC

after single step amplification and therefore a second round of amplification was performed using nested primers (Table 1).

For the analysis of mutations in PDGFRA exon 18, the following primer pairs used were 1) PDGFRA 181634F (residues 181634 through 181653 of SEQ ID NO: 19) and PDGFRA 181874R (residues 181844 through 181874 of SEQ ID NO: 19) or 2) PDGFRA 181752F (SNP exclusion) (residues 181752 through 181772 of SEQ ID NO: 19) and PDGFRA 181874R. The locations of these primers are indicated in Figure 7A, along with PDGFRA 181671F (residues 181671 through 181690 of SEQ ID NO: 19) and PDGFRA 181862R (residues 181842 through 181862 of SEQ ID NO: 19).

For the analysis of mutations in PDGFRA exon 12, the following primer pairs were used: 1) PDGFRA 170636F (residues 170636 through 170655 of SEQ ID NO: 19) and PDGFRA 170894R (residues 170876 through 170894 of SEQ ID NO: 19), and 2) PDGFRA 170658F (residues 170658 through 170677 of SEQ ID NO: 19) and PDGFRA 170866R (residues 170847 through 170866 of SEQ ID NO: 19).

Five to 20 µl aliquots of the final PCR reaction were screened for mutations on a Transgenomic WAVE HPLC system (D-HPLC; Transgenomic, Inc., Omaha, NE) by running at non-denaturing (50 °C) or partially denaturing temperature (61 °C). D-HPLC-detected mutations were confirmed by two methods: 1) re-amplification of the exon and repeat D-HPLC analysis on a different day; 2) bi-directional sequence analysis on an ABI 377 sequencer using the BigDye terminator kit (Applied Biosciences, Inc.). D-HPLC-detected mutations were confirmed by two methods: 1) re-amplification of the exon and repeat D-HPLC analysis on a different day; 2) bi-directional sequence analysis on an ABI 377 sequencer using the BigDye terminator kit (Applied Biosciences, Inc) (Corless *et al.*, *Am. J. Pathol.* 160, 1567, 2002).

Using primer pair 1, it was possible to reliably detect the D842V point mutation as well as the deletion and insertion mutations (Figures 3 and 4). However, there is a fairly common single nucleotide polymorphism (SNP) in the *PDGFRA* gene that is detected using these primer pairs and D-HPLC analysis. This SNP is C2472T (V824V) in *PDGFRA* cDNA (using numbering system of Genbank Accession No. XM_011186). To exclude this SNP, the mutation detection assay was further optimized by using primer pair 2. The forward primer of this set begins immediately 3' of the SNP and thus the resultant amplicon from this primer set does not contain the SNP. Using this primer pair, the D842V activating mutation can be reliably detected and differentiated from the C2472T (V824V) SNP (Figure 5).

To further verify the sequence of the PDGFRA exon 18 deletion mutations we cloned the amplification products into pCR®4-TOPO using the TOPO TA cloning kit (Invitrogen, version H) and the ligated plasmids were used to transform competent *E. coli* (OneShot TOP10, Invitrogen). Isolated plasmids were screened for the mutant exon insert by PCR and D-HPLC. Direct sequence analysis of cloned mutant DNA confirmed the presence of an in-frame exon 18 deletion in these cases.

Results

Activation of PDGFRA in GISTs

Using methods described above, RTK activation was assessed in three GISTs lacking apparent KIT oncoproteins. This was accomplished by immunoprecipitating with panRTK antibodies, and then immunoblotting with an antibody against phosphotyrosine (Figure 1). Normally, KIT is heavy phosphorylated in GISTs and is one of the dominant tyrosine phosphorylated protein (Figure 1).

By sequentially stripping and reprobing the membrane with additional antibodies, the predominant RTK phosphoprotein appeared to be PDGFRA. The possibility of a highly activated PDGFRA protein was then confirmed by immunoprecipitating PDGFRA, using a specific antibody to this protein. These studies revealed that the highly activated phosphoRTK comigrated with equally strongly phosphorylated PDGFRA (Figure 1). Further, these studies showed that KIT was inactive (nonphosphorylated) in the GISTs with strongly phosphorylated PDGFRA. Therefore, the studies revealed that PDGFRA is highly activated in a subset of GISTs that lack KIT activation, and – furthermore – PDGFRA is the predominant activated RTK, and indeed one of the predominant tyrosine phosphorylated proteins (Figure 1) in those GISTs.

Additional studies indicated that KIT and PDGFRA oncoproteins are typically alternative, rather than synergistic, mechanisms of transformation in GISTs. Therefore, PDGFRA activation and high-level PDGFRA expression can be found in GISTs that have reduced levels of KIT expression (Figure 2) and that lack KIT genomic oncogenic mutations.

Analysis of Genomic Mechanisms of PDGFRA activation in GISTs

The large amount of phosphorylated PDGFRA in these GISTs suggested the possibility of activating mutations in the *PDGFRA* gene. Clues to a possible location for such mutations came from comparisons with other related kinases. As mentioned above, mutation of KIT is common in GISTs (approximately 80-90% of cases); mutations also occur in seminoma (25% of cases), mastocytosis (95%+) and rarely in cases of Acute myeloid leukemia (AML) (Heinrich *et al.*, *J. Clin. Oncol.*, 20: 1692-1703, 2002; Rubin *et al.*, *Cancer Res*, 61: 8118-8121, 2001; Lux *et al.*, *Am.J.Pathol.*, 156: 791-795, 2000). KIT mutations in GIST are found most commonly in the juxtamembrane and extracellular domains, as well as the first portion of the tyrosine kinase domain, whereas mutations in mastocytosis and seminoma are found in the activation loop located in the second portion of the tyrosine kinase domain (Hirota *et al.*, *J. Pathol.*, 193: 505-510, 2001; Lasota *et al.*, *Am.J Pathol.*, 157: 1091-1095, 2000; Lux *et al.*, *Am.J.Pathol.*, 156: 791-795, 2000; Ma *et al.*, *Blood*, 99: 1741-1744, 2002; Beghini *et al.*, *Blood*, 95: 726-727, 2000; Tian *et al.*, *Am.J.Pathol.*, 154: 1643-1647, 1999; Longley *et al.*, *Nature Genetics*, 12: 312-314, 1996). Somatic mutation of *FLT3* has also been associated with certain human malignancies. Mutation of *FLT3* has been reported in approximately 20-40% of cases of AML and rarely in Acute Lymphoblastic Leukemia. In AML,

mutations of FLT3 are most commonly found in the juxtamembrane domain and less commonly in the activation loop (Abu-Duhier *et al.*, *Br. J Haematol.*, 113: 983-988, 2001; Kottaridis *et al.*, *Blood*, 98: 1752-1759, 2001; Meshinchi *et al.*, *Blood*, 97: 89-94, 2001; Yamamoto *et al.*, *Blood*, 97: 2434-2439, 2001).

Based on the homology of PDGFRA to KIT and FLT3, we hypothesized that mutation of the PDGFRA activation loop in a subset of GISTs might result in activation of tyrosine kinase activity. Thus, we developed a polymerase chain reaction (PCR) based assay to test for mutations of the PDGFRA activation loop (exon 18) (see Figure 7). Genomic DNA was purified from formalin fixed, paraffin embedded archival pathology specimens or fresh frozen tumor specimens that were obtained in accordance with the rules and regulations of both OHSU and the Portland VA Medical Center. Amplification of PDGFRA exon 18 was performed using primer sets described in the methods section below. Amplicons were analyzed using a Transgenomic WAVE HPLC instrument using both non-denaturing (50 °C) and partially denaturing temperatures (58 °C). Amplicons with abnormal HPLC elution profiles were directly sequenced.

Two different classes of PDGFRA activation loop mutations were identified in GISTs using this technique -- point mutation and small in-frame deletions (Figure 3). These amplicons have been directly sequenced and/or cloned into plasmids and the resultant clones sequenced. The most common mutation is a change of the conserved aspartic acid at position 842 of PDGFRA to valine (D842V). This aspartic acid is highly conserved in kinases related to PDGFRA. The homologous mutation D816V of KIT is observed in mastocytosis and seminoma, while the homologous D835V mutation of FLT3 is found in some cases of AML.

Two different in-frame deletions of PDGFRA exon 18 were identified in GISTs. The first is deletion of genomic nucleotides 53264-53275, which encode PDGRA amino residues 842-845 (DIMH). In this mutation the conserved aspartic residue at position 842 is substituted by the aspartic acid at position 846 that is immediately 3' of the deletion. The second deletion found to date is a deletion with insertion of a single cytosine at the 3' end of the deletion -- the result is deletion of residues 845-848 (HDSN) with generation of a novel proline residue that follows the normal methionine residue at position 844. Thus, these two deletions are partially overlapping. These deletions are novel; it is believed that they result in constitutive activation of the tyrosine kinase activity of PDGFRA. This is based on prior observations that in-frame deletions or insertion in the activation loop of the related FLT3 RTK are known to result in constitutive activation of tyrosine kinase activity (Abu-Duhier *et al.*, *Br. J Haematol.*, 113: 983-988, 2001); and our observations that PDGFRA is strongly activated in protein lysates from GIST tumors that harbor these PDGFRA mutations, but not in GISTs expressing wild-type PDGFRA (see Figures 1 and 2).

We have also found one GIST with an acquired mutation of exon 12 of PDGFRA, specifically insertion of GAGAGG at nucleotide position 1681 of PDGFRA. This mutation results in insertion of novel amino acid residues ER between amino acids 560 and 561. Based on analogy with similar length mutations in FLT3 and KIT, this inframe insertion would be predicted to result in

constitutive activation of PDGFRA kinase activity. We have also found a second example of an insertion/deletion mutation in exon 12 in a GIST: SPDGHE566-571R.

Table 1

Genotype	DNA sequence (top line) Translation (bottom line)
PDGFRA Wild type (Ac. No. XM_011186; SEQ ID NOs: 1 and 2)	2906* GGCCTGGCCAGAGACATCATGCATGATTCGAACTATGTG 838 G L A R D I M H D S N Y V
D842V (SEQ ID NOs: 3 and 4)	2906 GGCCTGGCCAGAGTCATCATGCATGATTCGAACTATGTG 838 G L A R V I M H D S N Y V
Deletion of DIMH842-845 (SEQ ID NOs: 5 and 6)	2906 GGCCTGGCCAGA----------GATTCGAACTATGTG 838 G L A R - - - - D S N Y V
Deletion of HSDN845-848P (SEQ ID NOs: 7 and 8)	2906 GGCCTGGCCAGAGACATCATGC---------CCTATGTG 838 G L A R D I M P Y V
PDGFRA Wild type	2060 GAAATTCGCTGGAGGGTCATTGAATCA 556 E I R W R V I E S
PDGFRA Insertion ER561-562 (SEQ ID NOs: 9 and 10)	2060 GAAATTCGCTGGAGGGAGAGGGTCATTGAATCA 556 E I R W R E R V I E S
PDGFRA Wild type	2081 GAATCAATCAGCCCGGATGGACATGAATATATT 563 E S I S P D G H E Y I
PDGFRA Deletion SPDGHE566-571R (SEQ ID NOs: 11 and 12)	2081 GAATCAATC--------------CCGTATATT 563 E S I - - - - - R Y I

*Numbering as in SEQ ID NO: 1 and SEQ ID NO: 2.

Table 2

Mutation	Cases (% total)
D842V	10 (24.4%)
Exon 18 Deletion	2 (4.9%)
Exon 12 Insertion/Deletion	2 (4.9%)
No mutation	27 (65.9%)
Total	**41 (100.0%)**

In our analysis of GISTs to date, we have found KIT mutation and PDGFRA mutation to be mutually exclusive. That is, PDGFRA mutations have only been found in GISTs without any detectable KIT mutation. Based on our studies to date, we believe that mutations of PDGFRA are found in approximately 34-35 % of KIT wild-type GISTs or 3-6% of all GISTs.

EXAMPLE 2: Other Activating PDGFRA Mutations

With the provision herein of the correlation between activating PDGFRA mutations and neoplastic disease, the isolation and identification of additional activating PDGFRA mutations is enabled. Any conventional method for the identification of genetic mutations in a population can be used to identify such additional mutations.

For instance, existing populations (*e.g.*, human populations) are assessed for the presence of neoplastic or tumorous cells, and individuals within the population are genotyped as relates to a PDGFRA sequence. These PDGFRA sequences are then compared to a reference PDGFRA

sequence, such as the alleles described herein, to determine the presence of one or more variant nucleotide positions. Once variant nucleotides are identified, statistical analysis of the population is used to determine whether these variants are correlated with neoplasm or tumorous growth or development.

By way of example, it is predicted that additional mutations will be identified at least in positions similar to those identified herein. SEQ ID NO: 26 shows a nucleic acid consensus sequence for the PDGFRA activating mutations discussed herein; the consensus polypeptide encoded by SEQ ID NO: 26 is shown in SEQ ID NO: 27. Explicitly contemplated herein are additional PDGFRA mutations and variant molecules that occur in the variable positions indicated in these consensus sequences, alone or in combination with one or more of the mutations described herein. Included are insertion and deletion mutations, such as examples provided herein, as well as point mutations.

EXAMPLE 3: Clinical Uses of PDGFRA Variants

To perform a diagnostic test for the presence or absence of a mutation in a PDGFRA sequence of an individual, a suitable genomic DNA-containing sample from a subject is obtained and the DNA extracted using conventional techniques. For instance, a blood sample, a buccal swab, a hair follicle preparation, or a nasal aspirate is used as a source of cells to provide the DNA sample; similarly, a surgical specimen, biopsy, or other biological sample containing genomic DNA could be used. It is particularly contemplated that tumor biopsies or tumor DNA found in plasma or other blood products can serve as a source. The extracted DNA is then subjected to amplification, for example according to standard procedures. The allele of the single base-pair mutation is determined by conventional methods including manual and automated fluorescent DNA sequencing, primer extension methods (Nikiforov, *et al.*, *Nucl Acids Res.* 22:4167-4175, 1994), oligonucleotide ligation assay (OLA) (Nickerson *et al.*, *Proc. Natl. Acad. Sci. USA* 87:8923-8927, 1990), allele-specific PCR methods (Rust *et al.*, *Nucl. Acids Res.* 6:3623-3629, 1993), RNase mismatch cleavage, single strand conformation polymorphism (SSCP), denaturing gradient gel electrophoresis (DGGE), Taq-Man™, oligonucleotide hybridization, and the like. Also, see the following U.S. Patents for descriptions of methods or applications of polymorphism analysis to disease prediction and/or diagnosis: 4,666,828 (RFLP for Huntington's); 4,801,531 (prediction of atherosclerosis); 5,110,920 (HLA typing); 5,268,267 (prediction of small cell carcinoma); and 5,387,506 (prediction of dysautonomia).

Examples of activating tyrosine kinase mutations are the PDGFRA D842V and V561D point mutations, the ER561-562 in frame insertion, and the DIMH842-845, HDSN845-848P, RD841-842KI, RVIES560-564, and SPDGHE566-571R in-frame deletions. In addition to these particular mutations, other mutations can be detected that may be associated with variable predisposition to development of a neoplastic disease or likelihood of having a tumor, and used in combination with the disclosed PDGFRA mutations, to predict the probability that a subject will develop neoplasia, or have a tumor with drug responsive tyrosine kinase activity.

The activating mutations of the present disclosure can be utilized for the detection of, and differentiation of, individuals who are homozygous and heterozygous for activating and/or drug responsive variants. The value of identifying individuals who carry an activating allele of PDGFRA (*i.e.*, individuals who are heterozygous or homozygous for an allele that contains the D842V or V561D point mutation, the ER561-562 in frame insertion, or one of the DIMH842-845, HDSN845-848P, RD841-842KI, RVIES560-564, or SPDGHE566-571R in-frame deletions, or any combination thereof, or another mutation in one or proximal to one of the variable regions indicated in SEQ ID NOs: 26 or 27) is that these individuals could then initiate customized therapies (such as specific drug therapies that inhibit the mutant, activated, PDGFRA), or undergo more aggressive treatment of the condition, and thereby beneficially alter its course.

EXAMPLE 4: Mutation Gene Probes and Markers

Sequences surrounding and overlapping single base-pair mutations and deletions and insertions in the PDGFRA gene can be useful for a number of gene mapping, targeting, and detection procedures. For example, genetic probes can be readily prepared for hybridization and detection of the D842V or the V561D point mutation, the ER561-562 in frame insertion, or one of the DIMH842-845, HDSN845-848P, RD841-842KI, RVIES560-564, or SPDGHE566-571R in-frame deletion mutations. As will be appreciated, probe sequences may be greater than about 12 or more oligonucleotides in length and possess sufficient complementarity to distinguish between the variant sequence and the wildtype, for instance, between the Valine (at amino acid residue 842 in the D842V activating allele) and Aspartic acid (in the wildtype allele). Similarly, sequences surrounding and overlapping any of the specifically disclosed mutations (or other mutations found in accordance with the present teachings, including those encompassed in or proximal to the variable regions indicated in SEQ ID NOs: 26 or 27), or longer sequences encompassing for instance the entire length of exon 18 of PDGFRA, or portions thereof, can be utilized in allele specific hybridization procedures. A similar approach can be adopted to detect other PDGFRA mutations.

Sequences surrounding and overlapping a PDGFRA mutation, or any portion or subset thereof that allows one to identify the mutations, are highly useful. Thus, another embodiment provides a genetic marker predictive of the one or more of the D842V or the V561D point mutation, the ER561-562 in frame insertion, or the DIMH842-845, HDSN845-848P, RD841-842KI, RVIES560-564, or SPDGHE566-571R in-frame deletions of PDGFRA, comprising a partial sequence of the human genome including at least about 10 contiguous nucleotide residues such as those shown in Table 1 or Table 3, and sequences complementary therewith.

EXAMPLE 5: Detecting Single Nucleotide Alterations

PDGFRA single nucleotide alterations, whether categorized as SNPs or new mutations (such as that giving rise to the D842V variant) can be detected by a variety of techniques. Clinically relevant PDGFRA single nucleotide alterations include those arising as somatic mutations – *i.e.*,

restricted to the neoplastic cells – as well as those that are present constitutionally in both normal and neoplastic cells in a given individual. The constitutional single nucleotide alterations can arise either from new germline mutations, or can be inherited from a parent who possesses a SNP or mutation in their own germline DNA. The techniques used in evaluating either somatic or germline single nucleotide alterations include allele-specific oligonucleotide hybridization (ASOH) (Stoneking *et al.*, *Am. J. Hum. Genet.* 48:370-382, 1991) which involves hybridization of probes to the sequence, stringent washing, and signal detection. Other new methods include techniques that incorporate more robust scoring of hybridization. Examples of these procedures include the ligation chain reaction (ASOH plus selective ligation and amplification), as disclosed in Wu and Wallace (*Genomics* 4:560-569, 1989); mini-sequencing (ASOH plus a single base extension) as discussed in Syvanen (*Meth. Mol. Biol.* 98:291-298, 1998); and the use of DNA chips (miniaturized ASOH with multiple oligonucleotide arrays) as disclosed in Lipshutz *et al.* (*BioTechniques* 19:442-447, 1995). Alternatively, ASOH with single- or dual- labeled probes can be merged with PCR, as in the 5'-exonuclease assay (Heid *et al.*, *Genome Res.* 6:986-994, 1996), or with molecular beacons (as in Tyagi and Kramer, *Nat. Biotechnol.* 14:303-308, 1996).

Another technique is dynamic allele-specific hybridization (DASH), which involves dynamic heating and coincident monitoring of DNA denaturation, as disclosed by Howell *et al.* (*Nat. Biotech.* 17:87-88, 1999). A target sequence is amplified by PCR in which one primer is biotinylated. The biotinylated product strand is bound to a streptavidin-coated microtiter plate well, and the non-biotinylated strand is rinsed away with alkali wash solution. An oligonucleotide probe, specific for one allele, is hybridized to the target at low temperature. This probe forms a duplex DNA region that interacts with a double strand-specific intercalating dye. When subsequently excited, the dye emits fluorescence proportional to the amount of double-stranded DNA (probe-target duplex) present. The sample is then steadily heated while fluorescence is continually monitored. A rapid fall in fluorescence indicates the denaturing temperature of the probe-target duplex. Using this technique, a single-base mismatch between the probe and target results in a significant lowering of melting temperature (T_m) that can be readily detected.

A variety of other techniques can be used to detect the mutations in DNA. Merely by way of example, see U.S. Patents No. 4,666,828; 4,801,531; 5,110,920; 5,268,267; 5,387,506; 5,691,153; 5,698,339; 5,736,330; 5,834,200; 5,922,542; and 5,998,137 for such methods.

EXAMPLE 6: Detection of PDGFRA Nucleic Acid Level(s)

Individuals carrying activating mutations in the PDGFRA gene, or having amplifications or heterozygous or homozygous deletions of the PDGFRA gene, may be detected at the DNA or RNA level with the use of a variety of techniques. The detection of point mutations, or SNPs, was discussed above; in the following example, techniques are provided for detecting the level of PDGFRA nucleic acid molecules in a sample.

For such diagnostic procedures, a biological sample of the subject (an animal, such as a mouse or a human), which biological sample contains either DNA or RNA derived from the subject, is assayed for a mutated, amplified or deleted PDGFRA encoding sequence, such as a genomic amplification of the PDGFRA gene or an over- or under-abundance of a PDGFRA mRNA. Suitable biological samples include samples containing genomic DNA or mRNA obtained from subject body cells, such as those present in peripheral blood, urine, saliva, tissue biopsy, surgical specimen, amniocentesis samples and autopsy material. The detection in the biological sample of a mutant PDGFRA gene, a mutant PDGFRA RNA, or an amplified or homozygously or heterozygously deleted PDGFRA gene, may be performed by a number of methodologies.

Gene dosage (copy number) can be important in disease states, and can influence mRNA and thereby protein level; it is therefore advantageous to determine the number of copies of PDGFRA nucleic acids in samples of tissue. Probes generated from the encoding sequence of PDGFRA (PDGFRA probes or primers) can be used to investigate and measure genomic dosage of the PDGFRA gene.

Appropriate techniques for measuring gene dosage are known in the art; see for instance, US Patent No. 5,569,753 ("Cancer Detection Probes") and Pinkel *et al.* (*Nat. Genet.* 20:207-211, 1998) ("High Resolution Analysis of DNA Copy Number Variation using Comparative Genomic Hybridization to Microarrays").

Determination of gene copy number in cells of a patient-derived sample using other techniques is known in the art. For example, PDGFRA amplification in immortalized cell lines as well as uncultured cells taken from a subject can be carried out using bicolor FISH or chromogenic *in situ* hybridization (CISH) analysis. FISH or CISH evaluations of PDGFRA amplification can be performed in various cell and tissue preparations that include, but are not limited to, venipuncture, biopsy, fine needle aspiration, and cell scraping. Such clinical materials can be analyzed in various forms, which include, but are not limited to, cytogenetic preparations; touch preparations from fresh or frozen biopsies; disaggregated cells from fresh, frozen or paraffin-embedded materials; histological sections from frozen or paraffin-embedded materials; and cytological preparations including cytospins and cell smears (Xiao *et al.*, *Am J Pathol;* Hsi *et al. Pathol.* 147:896-904; 1995; Davison *et al.*, *Am. J. Pathol.* 153:1401-1409; 1998. By way of example, interphase FISH analysis of immortalized cell lines can be carried out as previously described (Barlund *et al.*, *Genes Chromo. Cancer* 20:372-376, 1997). The hybridizations can be evaluated using a Zeiss fluorescence microscope. By way of example, approximately 20 non-overlapping nuclei with intact morphology based on DAPI counterstain are scored to determine the mean number of hybridization signals for each test and reference probe.

Likewise, FISH can be performed on tissue microarrays, as described in Kononen *et al.*, *Nat. Med.* 4:844-847, 1998. Briefly, consecutive sections of the array are deparaffinized, dehydrated in ethanol, denatured at 74° C for 5 minutes in 70% formamide/2 x SSC, and hybridized with test and reference probes. The specimens containing tight clusters of signals or >3-fold increase in the

number of test probe as compared to chromosome 17 centromere in at least 10% of the tumor cells may be considered as amplified. Microarrays using various tissues can be constructed as described in WO 99/44063 and WO 99/44062.

Overexpression of the PDGFRA gene can also be detected by measuring the cellular level of PDGFRA -specific mRNA. mRNA can be measured using techniques well known in the art, including for instance Northern analysis, RT-PCR and mRNA *in situ* hybridization.

EXAMPLE 7: Expression of PDGFRA Polypeptides

The expression and purification of proteins, such as the PDGFRA protein, can be performed using standard laboratory techniques. After expression, purified PDGFRA protein may be used for functional analyses, antibody production, diagnostics, and patient therapy. Furthermore, the DNA sequence of the PDGFRA cDNA can be manipulated in studies to understand the expression of the gene and the function of its product. Mutant forms of the human PDGFRA gene may be isolated based upon information contained herein, and may be studied in order to detect alteration in expression patterns in terms of relative quantities, tissue specificity and functional properties of the encoded mutant PDGFRA protein. Partial or full-length cDNA sequences, which encode for the subject protein, may be ligated into bacterial expression vectors. Methods for expressing large amounts of protein from a cloned gene introduced into *Escherichia coli* (*E. coli*) may be utilized for the purification, localization and functional analysis of proteins. For example, fusion proteins consisting of amino terminal peptides encoded by a portion of the *E. coli lacZ* or *trpE* gene linked to PDGFRA proteins may be used to prepare polyclonal and monoclonal antibodies against these proteins. Thereafter, these antibodies may be used to purify proteins by immunoaffinity chromatography, in diagnostic assays to quantitate the levels of protein and to localize proteins in tissues and individual cells by immunofluorescence.

Intact native protein may also be produced in *E. coli* in large amounts for functional studies. Methods and plasmid vectors for producing fusion proteins and intact native proteins in bacteria are described in Sambrook *et al.* (In *Molecular Cloning: A Laboratory Manual*, Ch. 17, CSHL, New York, 1989). Such fusion proteins may be made in large amounts, are easy to purify, and can be used to elicit antibody response. Native proteins can be produced in bacteria by placing a strong, regulated promoter and an efficient ribosome-binding site upstream of the cloned gene. If low levels of protein are produced, additional steps may be taken to increase protein production; if high levels of protein are produced, purification is relatively easy. Suitable methods are presented in Sambrook *et al.* (In *Molecular Cloning: A Laboratory Manual*, CSHL, New York, 1989) and are well known in the art. Often, proteins expressed at high levels are found in insoluble inclusion bodies. Methods for extracting proteins from these aggregates are described by Sambrook *et al.* (In *Molecular Cloning: A Laboratory Manual*, Ch. 17, CSHL, New York, 1989). Vector systems suitable for the expression of *lacZ* fusion genes include the pUR series of vectors (Ruther and Muller-Hill, *EMBO J.* 2:1791, 1983), pEX1-3 (Stanley and Luzio, *EMBO J.* 3:1429, 1984) and pMR100 (Gray *et al., Proc. Natl. Acad. Sci.*

USA 79:6598, 1982). Vectors suitable for the production of intact native proteins include pKC30 (Shimatake and Rosenberg, *Nature* 292:128, 1981), pKK177-3 (Amann and Brosius, *Gene* 40:183, 1985) and pET-3 (Studiar and Moffatt, *J. Mol. Biol.* 189:113, 1986). Fusion proteins may be isolated from protein gels, lyophilized, ground into a powder and used as an antigen. The DNA sequence can also be transferred from its existing context to other cloning vehicles, such as other plasmids, bacteriophages, cosmids, animal viruses and yeast artificial chromosomes (YACs) (Burke *et al.*, *Science* 236:806-812, 1987). These vectors may then be introduced into a variety of hosts including somatic cells, and simple or complex organisms, such as bacteria, fungi (Timberlake and Marshall, *Science* 244:1313-1317, 1989), invertebrates, plants (Gasser and Fraley, *Science* 244:1293, 1989), and animals (Pursel *et al.*, *Science* 244:1281-1288, 1989), which cell or organisms are rendered transgenic by the introduction of the heterologous PDGFRA cDNA.

For expression in mammalian cells, the cDNA sequence may be ligated to heterologous promoters, such as the simian virus (SV) 40 promoter in the pSV2 vector (Mulligan and Berg, *Proc. Natl. Acad. Sci. USA* 78:2072-2076, 1981), and introduced into cells, such as monkey COS-1 cells (Gluzman, *Cell* 23:175-182, 1981), to achieve transient or long-term expression. The stable integration of the chimeric gene construct may be maintained in mammalian cells by biochemical selection, such as neomycin (Southern and Berg, *J. Mol. Appl. Genet.* 1:327-341, 1982) and mycophenolic acid (Mulligan and Berg, *Proc. Natl. Acad. Sci. USA* 78:2072-2076, 1981).

DNA sequences can be manipulated with standard procedures such as restriction enzyme digestion, fill-in with DNA polymerase, deletion by exonuclease, extension by terminal deoxynucleotide transferase, ligation of synthetic or cloned DNA sequences, site-directed sequence-alteration via single-stranded bacteriophage intermediate or with the use of specific oligonucleotides in combination with PCR or other *in vitro* amplification.

The cDNA sequence (or portions derived from it) or a mini gene (a cDNA with an intron and its own promoter) may be introduced into eukaryotic expression vectors by conventional techniques. These vectors are designed to permit the transcription of the cDNA in eukaryotic cells by providing regulatory sequences that initiate and enhance the transcription of the cDNA and ensure its proper splicing and polyadenylation. Vectors containing the promoter and enhancer regions of the SV40 or long terminal repeat (LTR) of the Rous Sarcoma virus and polyadenylation and splicing signal from SV40 are readily available (Mulligan *et al.*, *Proc. Natl. Acad. Sci. USA* 78:1078-2076, 1981; Gorman *et al.*, *Proc. Natl. Acad. Sci USA* 78:6777-6781, 1982). The level of expression of the cDNA can be manipulated with this type of vector, either by using promoters that have different activities (for example, the baculovirus pAC373 can express cDNAs at high levels in *S. frugiperda* cells (Summers and Smith, In *Genetically Altered Viruses and the Environment*, Fields *et al.* (Eds.) 22:319-328, CSHL Press, Cold Spring Harbor, New York, 1985) or by using vectors that contain promoters amenable to modulation, for example, the glucocorticoid-responsive promoter from the mouse mammary tumor virus (Lee *et al.*, *Nature* 294:228, 1982). The expression of the cDNA can be monitored in the recipient cells 24 to 72 hours after introduction (transient expression).

In addition, some vectors contain selectable markers such as the *gpt* (Mulligan and Berg, *Proc. Natl. Acad. Sci. USA* 78:2072-2076, 1981) or *neo* (Southern and Berg, *J. Mol. Appl. Genet.* 1:327-341, 1982) bacterial genes. These selectable markers permit selection of transfected cells that exhibit stable, long-term expression of the vectors (and therefore the cDNA). The vectors can be maintained in the cells as episomal, freely replicating entities by using regulatory elements of viruses such as papilloma (Sarver *et al., Mol. Cell Biol.* 1:486, 1981) or Epstein-Barr (Sugden *et al., Mol. Cell Biol.* 5:410, 1985). Alternatively, one can also produce cell lines that have integrated the vector into genomic DNA. Both of these types of cell lines produce the gene product on a continuous basis. One can also produce cell lines that have amplified the number of copies of the vector (and therefore of the cDNA as well) to create cell lines that can produce high levels of the gene product (Alt *et al., J. Biol. Chem.* 253:1357, 1978).

The transfer of DNA into eukaryotic, in particular human or other mammalian cells, is now a conventional technique. The vectors are introduced into the recipient cells as pure DNA (transfection) by, for example, precipitation with calcium phosphate (Graham and vander Eb, *Virology* 52:466, 1973) or strontium phosphate (Brash *et al., Mol. Cell Biol.* 7:2013, 1987), electroporation (Neumann *et al., EMBO J* 1:841, 1982), lipofection (Felgner *et al., Proc. Natl. Acad. Sci USA* 84:7413, 1987), DEAE dextran (McCuthan *et al., J. Natl. Cancer Inst.* 41:351, 1968), microinjection (Mueller *et al., Cell* 15:579, 1978), protoplast fusion (Schafner, *Proc. Natl. Acad. Sci. USA* 77:2163-2167, 1980), or pellet guns (Klein *et al., Nature* 327:70, 1987). Alternatively, the cDNA, or fragments thereof, can be introduced by infection with virus vectors. Systems are developed that use, for example, retroviruses (Bernstein *et al., Gen. Engr'g* 7:235, 1985), adenoviruses (Ahmad *et al., J. Virol.* 57:267, 1986), or Herpes virus (Spaete *et al., Cell* 30:295, 1982). Tyrosine kinase encoding sequences can also be delivered to target cells *in vitro* via non-infectious systems, for instance liposomes.

These eukaryotic expression systems can be used for studies of PDGFRA encoding nucleic acids and mutant forms of these molecules, the PDGFRA protein and mutant forms of this protein. Such uses include, for example, the identification of regulatory elements located in the 5' region of the PDGFRA gene on genomic clones that can be isolated from human genomic DNA libraries using the information contained in the present disclosure. The eukaryotic expression systems may also be used to study the function of the normal complete protein, specific portions of the protein, or of naturally occurring or artificially produced mutant proteins.

Using the above techniques, the expression vectors containing the PDGFRA gene sequence or cDNA, or fragments or variants or mutants thereof, can be introduced into human cells, mammalian cells from other species or non-mammalian cells as desired. The choice of cell is determined by the purpose of the treatment. For example, monkey COS cells (Gluzman, *Cell* 23:175-182, 1981) that produce high levels of the SV40 T antigen and permit the replication of vectors containing the SV40 origin of replication may be used. Similarly, Chinese hamster ovary (CHO), mouse NIH 3T3 fibroblasts or human fibroblasts or lymphoblasts may be used.

The present disclosure thus encompasses recombinant vectors that comprise all or part of the PDGFRA gene or cDNA sequences, for expression in a suitable host. The PDGFRA DNA is operatively linked in the vector to an expression control sequence in the recombinant DNA molecule so that the PDGFRA polypeptide can be expressed. The expression control sequence may be selected from the group consisting of sequences that control the expression of genes of prokaryotic or eukaryotic cells and their viruses and combinations thereof. The expression control sequence may be specifically selected from the group consisting of the *lac* system, the *trp* system, the *tac* system, the *trc* system, major operator and promoter regions of phage lambda, the control region of fd coat protein, the early and late promoters of SV40, promoters derived from polyoma, adenovirus, retrovirus, baculovirus and simian virus, the promoter for 3-phosphoglycerate kinase, the promoters of yeast acid phosphatase, the promoter of the yeast alpha-mating factors and combinations thereof.

The host cell, which may be transfected with the vector of this disclosure, may be selected from the group consisting of *E. coli*, *Pseudomonas*, *Bacillus subtilis*, *Bacillus stearothermophilus* or other bacilli; other bacteria; yeast; fungi; insect; mouse or other animal; or plant hosts; or human tissue cells.

It is appreciated that for mutant or variant PDGFRA DNA sequences, similar systems are employed to express and produce the mutant product. In addition, fragments of the PDGFRA protein can be expressed essentially as detailed above. Such fragments include individual PDGFRA protein domains or sub-domains, as well as shorter fragments such as peptides. PDGFRA protein fragments having therapeutic properties may be expressed in this manner also.

EXAMPLE 8: Production of PDGFRA Protein Specific Binding Agents

Monoclonal or polyclonal antibodies may be produced to either the normal PDGFRA protein or mutant forms of this protein, for instance particular portions that contain a mutation and therefore may provide a distinguishing epitope. Optimally, antibodies raised against these proteins or peptides would specifically detect the protein or peptide with which the antibodies are generated. That is, an antibody generated to the PDGFRA protein or a fragment thereof would recognize and bind the PDGFRA protein and would not substantially recognize or bind to other proteins found in human cells. In some embodiments, an antibody is specific for (or measurably preferentially binds to) an epitope in a variant protein versus the wildtype protein, or vice versa, as discussed more fully herein.

The determination that an antibody specifically detects the PDGFRA protein is made by any one of a number of standard immunoassay methods; for instance, the western blotting technique (Sambrook *et al.*, In *Molecular Cloning: A Laboratory Manual*, CSHL, New York, 1989). To determine that a given antibody preparation (such as one produced in a mouse) specifically detects the PDGFRA protein by western blotting, total cellular protein is extracted from human cells (for example, lymphocytes) and electrophoresed on a sodium dodecyl sulfate-polyacrylamide gel. The proteins are then transferred to a membrane (for example, nitrocellulose) by western blotting, and the

antibody preparation is incubated with the membrane. After washing the membrane to remove non-specifically bound antibodies, the presence of specifically bound antibodies is detected by the use of an anti-mouse antibody conjugated to an enzyme such as alkaline phosphatase. Application of an alkaline phosphatase substrate 5-bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazolium results in the production of a dense blue compound by immunolocalized alkaline phosphatase. Antibodies that specifically detect the PDGFRA protein will, by this technique, be shown to bind to the PDGFRA protein band (which will be localized at a given position on the gel determined by its molecular weight). Non-specific binding of the antibody to other proteins may occur and may be detectable as a weak signal on the Western blot. The non-specific nature of this binding will be recognized by one skilled in the art by the weak signal obtained on the Western blot relative to the strong primary signal arising from the specific antibody-PDGFRA protein binding.

Substantially pure PDGFRA protein or protein fragment (peptide) suitable for use as an immunogen may be isolated from the transfected or transformed cells as described above. Concentration of protein or peptide in the final preparation is adjusted, for example, by concentration on an Amicon filter device, to the level of a few micrograms per milliliter. Monoclonal or polyclonal antibody to the protein can then be prepared as follows:

A. Monoclonal Antibody Production by Hybridoma Fusion

Monoclonal antibody to epitopes of the PDGFRA protein identified and isolated as described can be prepared from murine hybridomas according to the classical method of Kohler and Milstein (*Nature* 256:495-497, 1975) or derivative methods thereof. Briefly, a mouse is repetitively inoculated with a few micrograms of the selected protein over a period of a few weeks. The mouse is then sacrificed, and the antibody-producing cells of the spleen isolated. The spleen cells are fused by means of polyethylene glycol with mouse myeloma cells, and the excess un-fused cells destroyed by growth of the system on selective media comprising aminopterin (HAT media). The successfully fused cells are diluted and aliquots of the dilution placed in wells of a microtiter plate where growth of the culture is continued. Antibody-producing clones are identified by detection of antibody in the supernatant fluid of the wells by immunoassay procedures, such as ELISA, as originally described by Engvall (*Meth. Enzymol.* 70:419-439, 1980), and derivative methods thereof. Selected positive clones can be expanded and their monoclonal antibody product harvested for use. Detailed procedures for monoclonal antibody production are described in Harlow and Lane (*Antibodies, A Laboratory Manual*, CSHL, New York, 1988).

B. Polyclonal Antibody Production by Immunization

Polyclonal antiserum containing antibodies to heterogeneous epitopes of a single protein can be prepared by immunizing suitable animals with the expressed protein (Example 7), which can be unmodified or modified to enhance immunogenicity. Effective polyclonal antibody production is affected by many factors related both to the antigen and the host species. For example, small molecules tend to be less immunogenic than others and may require the use of carriers and adjuvant.

Also, host animals vary in response to site of inoculations and dose, with either inadequate or excessive doses of antigen resulting in low titer antisera. Small doses (ng level) of antigen administered at multiple intradermal sites appear to be most reliable. An effective immunization protocol for rabbits can be found in Vaitukaitis *et al.* (*J. Clin. Endocrinol. Metab.* 33:988-991, 1971).

Booster injections can be given at regular intervals, and antiserum harvested when antibody titer thereof, as determined semi-quantitatively, for example, by double immunodiffusion in agar against known concentrations of the antigen, begins to fall. See, for example, Ouchterlony *et al.* (In *Handbook of Experimental Immunology*, Wier, D. (ed.) chapter 19. Blackwell, 1973). Plateau concentration of antibody is usually in the range of about 0.1 to 0.2 mg/ml of serum (about 12 μM). Affinity of the antisera for the antigen is determined by preparing competitive binding curves, as described, for example, by Fisher (*Manual of Clinical Immunology*, Ch. 42, 1980).

C. Antibodies Raised against Synthetic Peptides

A third approach to raising antibodies against the PDGFRA protein or peptides is to use one or more synthetic peptides synthesized on a commercially available peptide synthesizer based upon the predicted amino acid sequence of the PDGFRA protein or peptide. Polyclonal antibodies can be generated by injecting these peptides into, for instance, rabbits or mice.

D. Antibodies Raised by Injection of PDGFRA Encoding Sequence

Antibodies may be raised against PDGFRA proteins and peptides by subcutaneous injection of a DNA vector that expresses the desired protein or peptide, or a fragment thereof, into laboratory animals, such as mice. Delivery of the recombinant vector into the animals may be achieved using a hand-held form of the Biolistic system (Sanford *et al., Particulate Sci. Technol.* 5:27-37, 1987) as described by Tang *et al.* (*Nature* 356:152-154, 1992). Expression vectors suitable for this purpose may include those that express the PDGFRA encoding sequence under the transcriptional control of either the human β-actin promoter or the cytomegalovirus (CMV) promoter.

Antibody preparations prepared according to these protocols are useful in quantitative immunoassays which determine concentrations of antigen-bearing substances in biological samples; they are also used semi-quantitatively or qualitatively to identify the presence of antigen in a biological sample; or for immunolocalization of the PDGFRA protein.

In addition, antibodies to PDGFRA are commercially available. See, for instance, rabbit anti-PDGFRA, catalog no. sc-338, from Santa Cruz Biotechnology Inc. (Santa Cruz, CA) and rabbit ant-PDGFR, catalog no. 6495, from Upstate Biotechnology (Waltham, MA).

For administration to human patients, antibodies, *e.g.*, PDGFRA-specific monoclonal antibodies, can be humanized by methods known in the art. Antibodies with a desired binding specificity can be commercially humanized (Scotgene, Scotland, UK; Oxford Molecular, Palo Alto, CA).

E. Antibodies Specific for Mutant PDGFRA

With the provision of several activating variant PDGFRA proteins, the production of antibodies that specifically recognize these proteins (and peptides derived therefrom) is enabled. In

particular, production of antibodies (and fragments and engineered versions thereof) that recognize at least one PDGFRA variant with a higher affinity than they recognize wild type PDGFRA is beneficial, as the resultant antibodies can be used in diagnosis and treatment, as well as in study and examination of the PDGFRA proteins themselves.

In particular embodiments, it is beneficial to generate antibodies from a peptide taken from a mutation or variation-specific region of the PDGFRA protein. By way of example, such regions include a portion or all of exon 18 of PDGFRA, or a portion or all of exon 12. More particularly, it is beneficial to raise antibodies against peptides of four or more contiguous amino acids that overlap the mutations identified in SEQ ID NO: 4, 6, 8, or 25, and particularly which comprise at least four contiguous amino acids including the residue(s) shown in position(s) 842 of SEQ ID NO: 4, positions 841 and 842 of SEQ ID NO: 6, positions 846 and 847 of SEQ ID NO: 8, or positions 841 and 842 of SEQ ID NO: 25.

Similarly, it is beneficial to raise antibodies against peptides of 4 or more contiguous amino acids that overlap the mutations identified in SEQ ID NO: 10, 12, 21, or 23, and particularly which comprise at least four contiguous amino acids including the residue(s) shown in position(s) 561 and 562 of SEQ ID NO: 10 positions 565 and 566 of SEQ ID NO: 12, position 561 of SEQ ID NO: 21, or positions 559 and 560 of SEQ ID NO: 23.

Longer peptides also can be used, and in some instances will produce a stronger or more reliable immunogenic response. Thus, it is contemplated in some embodiments that more than four amino acids are used to elicit the immune response, for instance, at least 5, at least 6, at least 8, at least 10, at least 12, at least 15, at least 18, at least 20, at least 25, or more, such as 30, 40, 50, or even longer peptides. Also, it will be understood by those of ordinary skill that it is beneficial in some instances to include adjuvants and other immune response enhancers, including passenger peptides or proteins, when using peptides to induce an immune response for production of antibodies.

Embodiments are not limited to antibodies that recognize epitopes containing the actual mutation identified in each variant. Instead, it is contemplated that variant-specific antibodies also may each recognize an epitope located anywhere throughout the PDGFRA variant molecule, which epitopes are changed in conformation and/or availability because of the activating mutation. Antibodies directed to any of these variant-specific epitopes are also encompassed herein.

By way of example, the following references provide descriptions of methods for making antibodies specific to mutant proteins: Hills *et al.*, (*Int. J. Cancer*, 63: 537-543, 1995); Reiter & Maihle (*Nucleic Acids Res.*, 24: 4050-4056, 1996); Okamoto *et al.* (*Br. J. Cancer*, 73: 1366-1372, 1996); Nakayashiki *et al.*, (*Jpn. J. Cancer Res.*, 91: 1035-1043, 2000); Gannon *et al.* (*EMBO J.*, 9: 1595-1602, 1990); Wong *et al.* (*Cancer Res.*, 46: 6029-6033, 1986); and Carney *et al.* (*J. Cell Biochem.*, 32: 207-214, 1986). Similar methods can be employed to generate antibodies specific to specific PDGFRA variants.

EXAMPLE 9: Protein-Based Diagnosis

An alternative method of diagnosing PDGFRA mutation, gene amplification, or deletion as well as abnormal PDGFRA expression, is to quantitate the level of PDGFRA protein, and/or to evaluate activation (phosphorylation) of PDGFRA in the cells of an individual. The oncogenic, activating mutations disclosed herein result in constitutive PDGFRA activation as manifested by PDGFRA tyrosine phosphorylation. Therefore, antibodies specific for phosphotyrosine-containing PDGFRA epitopes can be used to routinely detect such mutant, activated, PDGFRA proteins in any mammalian cell type. Such evaluations can be performed, for example, in lysates prepared from cells, in fresh or frozen cells, in cells that have been smeared or touched on glass slides and then either fixed and/or dried, or in cells that have been fixed, embedded (*e.g.*, in paraffin), and then prepared as histological sections on glass slides. This diagnostic tool would also be useful for detecting reduced levels of the PDGFRA protein that result from, for example, mutations in the promoter regions of the PDGFRA gene or mutations within the coding region of the gene that produced truncated, non-functional or unstable polypeptides, as well as from deletions of a portion of or the entire PDGFRA gene. Alternatively, amplification of a PDGFRA-encoding sequence may be detected as an increase in the expression level of PDGFRA protein. Such an increase in protein expression may also be a result of an up-regulating mutation in the promoter region or other regulatory or coding sequence within the PDGFRA gene, or by virtue of a point mutation within the PDGFRA coding sequence, which protects the PDGFRA protein from degradation.

Localization and/or coordination of PDGFRA expression (temporally or spatially) can also be examined using known techniques, such as isolation and comparison of PDGFRA from subcellular fractions, including specific organelles, or from specific cell or tissue types, or at specific time points after an experimental manipulation. Demonstration of reduced or increased PDGFRA protein levels, in comparison to such expression in a control cell (*e.g.*, normal, as in taken from a subject not suffering from a neoplastic disease, such as cancer), would be an alternative or supplemental approach to the direct determination of PDGFRA gene deletion, amplification or mutation status by the methods outlined above and equivalents.

The availability of antibodies specific to the PDGFRA protein will facilitate the detection and quantitation of cellular PDGFRA by one of a number of immunoassay methods which are well known in the art and are presented in Harlow and Lane (*Antibodies, A Laboratory Manual*, CSHL, New York, 1988). Methods of constructing such antibodies are discussed above, in Example 8.

Any standard immunoassay format (*e.g.*, ELISA, western blot, or RIA assay) can be used to measure PDGFRA polypeptide or protein levels, and to compare these with PDGFRA expression levels in control, reference, cell populations. Altered PDGFRA polypeptide expression may be indicative of an abnormal biological condition related to unregulated cell growth or proliferation, in particular a neoplasm, and/or a predilection to development of neoplastic disease. Immunohistochemical techniques may also be utilized for PDGFRA polypeptide or protein detection. For example, a tissue sample may be obtained from a subject, and a section stained for the presence

of PDGFRA using a PDGFRA specific binding agent (*e.g.*, anti-PDGFRA antibody) and any standard detection system (*e.g.*, one which includes a secondary antibody conjugated to horseradish peroxidase). General guidance regarding such techniques can be found in, *e.g.*, Bancroft and Stevens (*Theory and Practice of Histological Techniques*, Churchill Livingstone, 1982) and Ausubel *et al.* (*Current Protocols in Molecular Biology*, John Wiley & Sons, New York, 1998).

For the purposes of quantitating a PDGFRA protein, a biological sample of the subject (which can be any animal, for instance a mouse or a human), which sample includes cellular proteins, is required. Such a biological sample may be obtained from body cells, such as those present in a tissue biopsy, surgical specimens, or autopsy material. In particular embodiments biological samples may be obtained from peripheral blood sample, urine, saliva, amniocentesis samples, and so forth. Quantitation of PDGFRA protein can be achieved by immunoassay and compared to levels of the protein found in control cells (*e.g.*, healthy, non-neoplastic cells of the same lineage or type as those under evaluation, or from a patient known not to have a neoplastic disease). Detection of tyrosine phosphorylated PDGFRA (using an antibody, *i.e.* a phospho-specific antibody, that detects such forms and does not detect non-phosphorylated PDGFRA) could be taken as an indication of a PDGFRA protein containing an activating mutation. Detection of phosphorylated PDGFRA could also indicate activation by other mechanisms, such as overexpression of PDGFRA by genomic amplification, or over-expression of PDGFRA ligands, *e.g.* PDGF-A. A significant (*e.g.*, 10% or greater) reduction in the amount of PDGFRA protein in the cells of a subject compared to the amount of PDGFRA protein found in normal human cells could be taken as an indication that the subject may have deletions or mutations in the PDGFRA gene, whereas a significant (*e.g.*, 10% or greater) increase would indicate that a duplication (amplification), or mutation that increases the stability of the PDGFRA protein or mRNA, may have occurred. Deletion, mutation, and/or amplification within the PDGFRA encoding sequence, and substantial under- or over-expression of PDGFRA protein, may be indicative of neoplastic disease (such as a tumor) and/or a predilection to develop neoplastic disease.

EXAMPLE 10: Differentiation of Individuals Homozygous versus Heterozygous for Activating Mutation(s)

Though it is believed that the activating variants described herein are the result of sporadic mutations rather than germline mutations, it may sometimes be beneficial to determine whether a subject is homozygous or heterozygous for the mutation.

By way of example, the oligonucleotide ligation assay (OLA), as described at Nickerson *et al.* (*Proc. Natl. Acad. Sci. USA* 87:8923-8927, 1990), allows the differentiation between individuals who are homozygous versus heterozygous for the D842V or the V561D point mutation, the ER561-562 in frame insertion, or the DIMH842-845, HDSN845-848P, RD841-842KI, RVIES560-564, or SPDGHE566-571R in-frame deletions. This feature allows one to rapidly and easily determine whether an individual is homozygous for at least one tyrosine kinase activating mutation, which

condition is linked to a relatively high predisposition to developing neoplastic disease and/or an increased likelihood of having a tumor. Alternatively, OLA can be used to determine whether a subject is homozygous for either of these mutations.

As an example of the OLA assay, when carried out in microtiter plates, one well is used for the determination of the presence of the PDGFRA allele that contains a T at nucleotide position 2919 (numbering from SEQ ID NO: 1) and a second well is used for the determination of the presence of the PDGFRA allele that contains an A at that nucleotide position in the wildtype sequence. Thus, the results for an individual who is heterozygous for the mutation will show a signal in each of the A and T wells.

EXAMPLE 11: Suppression of PDGFRA Expression

A reduction of PDGFRA protein expression in a transgenic cell may be obtained by introducing into cells an antisense construct based on the PDGFRA encoding sequence, including the human PDGFRA cDNA or genomic sequence (SEQ ID NOs: 1 and 19, respectively) or flanking regions thereof. For antisense suppression, a nucleotide sequence from a PDGFRA encoding sequence, *e.g.* all or a portion of the PDGFRA cDNA or gene, is arranged in reverse orientation relative to the promoter sequence in the transformation vector. Other aspects of the vector may be chosen as discussed above (Example 7).

The introduced sequence need not be the full-length human PDGFRA cDNA or gene or reverse complement thereof, and need not be exactly homologous to the equivalent sequence found in the cell type to be transformed. Generally, however, where the introduced sequence is of shorter length, a higher degree of homology to the native PDGFRA sequence will be needed for effective antisense suppression. The introduced antisense sequence in the vector may be at least 30 nucleotides in length, and improved antisense suppression will typically be observed as the length of the antisense sequence increases. The length of the antisense sequence in the vector advantageously may be greater than 100 nucleotides. For suppression of the PDGFRA gene itself, transcription of an antisense construct results in the production of RNA molecules that are the reverse complement of mRNA molecules transcribed from the endogenous PDGFRA gene in the cell.

Although the exact mechanism by which antisense RNA molecules interfere with gene expression has not been elucidated, it is believed that antisense RNA molecules bind to the endogenous mRNA molecules and thereby inhibit translation of the endogenous mRNA.

Suppression of endogenous PDGFRA expression can also be achieved using ribozymes. Ribozymes are synthetic RNA molecules that possess highly specific endoribonuclease activity. The production and use of ribozymes are disclosed in U.S. Patent No. 4,987,071 to Cech and U.S. Patent No. 5,543,508 to Haselhoff. The inclusion of ribozyme sequences within antisense RNAs may be used to confer RNA cleaving activity on the antisense RNA, such that endogenous mRNA molecules that bind to the antisense RNA are cleaved, which in turn leads to an enhanced antisense inhibition of endogenous gene expression.

Expression of PDGFRA can also be reduced using small inhibitory RNAs, for instance using techniques similar to those described previously (see, *e.g.*, Tuschl *et al.*, *Genes Dev* 13, 3191-3197, 1999; Caplen *et al.*, *Proc. Nat.l Acad. Sci. U. S. A.* 98, 9742-9747, 2001; and Elbashir *et al.*, *Nature* 411, 494-498, 2001).

Finally, dominant negative mutant forms of PDGFRA may be used to block endogenous PDGFRA activity.

EXAMPLE 12: PDGFRA Gene Therapy

Gene therapy approaches for combating activating mutations in PDGFRA, or reducing the risk of developing neoplastic disease such as cancer, in subjects are now made possible by the present disclosure.

Retroviruses have been considered a preferred vector for experiments in gene therapy, with a high efficiency of infection and stable integration and expression (Orkin *et al.*, *Prog. Med. Genet.* 7:130-142, 1988). The full-length PDGFRA gene or cDNA can be cloned into a retroviral vector and driven from either its endogenous promoter or from the retroviral LTR (long terminal repeat). Other viral transfection systems may also be utilized for this type of approach, including adenovirus, adeno-associated virus (AAV) (McLaughlin *et al.*, *J. Virol.* 62:1963-1973, 1988), *Vaccinia* virus (Moss *et al.*, *Annu. Rev. Immunol.* 5:305-324, 1987), Bovine Papilloma virus (Rasmussen *et al.*, *Methods Enzymol.* 139:642-654, 1987) or members of the herpesvirus group such as Epstein-Barr virus (Margolskee *et al.*, *Mol. Cell. Biol.* 8:2837-2847, 1988).

Recent developments in gene therapy techniques include the use of RNA-DNA hybrid oligonucleotides, as described by Cole-Strauss *et al.* (*Science* 273:1386-1389, 1996). This technique may allow for site-specific integration of cloned sequences, thereby permitting accurately targeted gene replacement.

In addition to delivery of a PDGFRA encoding sequence to cells using viral vectors, it is possible to use non-infectious methods of delivery. For instance, lipidic and liposome-mediated gene delivery has recently been used successfully for transfection with various genes (for reviews, see Templeton and Lasic, *Mol. Biotechnol.* 11:175-180, 1999; Lee and Huang, *Crit. Rev. Ther. Drug Carrier Syst.* 14:173-206; and Cooper, *Semin. Oncol.* 23:172-187, 1996). For instance, cationic liposomes have been analyzed for their ability to transfect monocytic leukemia cells, and shown to be a viable alternative to using viral vectors (de Lima *et al.*, *Mol. Membr. Biol.* 16:103-109, 1999). Such cationic liposomes can also be targeted to specific cells through the inclusion of, for instance, monoclonal antibodies or other appropriate targeting ligands (Kao *et al.*, *Cancer Gene Ther.* 3:250-256, 1996).

To reduce the level of PDGFRA expression, gene therapy can be carried out using antisense or other suppressive constructs, the construction of which is discussed above (Example 11).

EXAMPLE 13: Kits

Kits are provided which contain the necessary reagents for determining the presence or absence of mutation(s) in a PDGFRA-encoding sequence, such as probes or primers specific for the PDGFRA gene or a highly variable region of this gene, such as those regions indicated in SEQ ID NO: 26. Such kits can be used with the methods described herein to determine whether a subject is predisposed to neoplastic disease or tumor development, or whether the subject is expected to respond to one or another therapy, such as a particular tyrosine kinase inhibitory compound.

The provided kits may also include written instructions. The instructions can provide calibration curves or charts to compare with the determined (*e.g.*, experimentally measured) values. Kits are also provided to determine elevated or depressed expression of mRNA (*i.e.*, containing probes) or PDGFRA protein (*i.e.*, containing antibodies or other PDGFRA-protein specific binding agents).

A. Kits for Amplification of PDGFRA Sequences

Oligonucleotide probes and primers, including those disclosed herein, can be supplied in the form of a kit for use in detection of a predisposition to neoplastic disease or tumor formation in a subject. In such a kit, an appropriate amount of one or more of the oligonucleotide primers is provided in one or more containers. The oligonucleotide primers may be provided suspended in an aqueous solution or as a freeze-dried or lyophilized powder, for instance. The container(s) in which the oligonucleotide(s) are supplied can be any conventional container that is capable of holding the supplied form, for instance, microfuge tubes, ampoules, or bottles. In some applications, pairs of primers may be provided in pre-measured single use amounts in individual, typically disposable, tubes or equivalent containers. With such an arrangement, the sample to be tested for the presence of a PDGFRA mutation can be added to the individual tubes and amplification carried out directly.

The amount of each oligonucleotide primer supplied in the kit can be any appropriate amount, depending for instance on the market to which the product is directed. For instance, if the kit is adapted for research or clinical use, the amount of each oligonucleotide primer provided would likely be an amount sufficient to prime several PCR amplification reactions. Those of ordinary skill in the art know the amount of oligonucleotide primer that is appropriate for use in a single amplification reaction. General guidelines may for instance be found in Innis *et al.* (PCR Protocols, A Guide to Methods and Applications, Academic Press, Inc., San Diego, CA, 1990), Sambrook *et al.* (In Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, New York, 1989), and Ausubel *et al.* (In Current Protocols in Molecular Biology, Greene Publ. Assoc. and Wiley-Intersciences, 1992).

A kit may include more than two primers, in order to facilitate the *in vitro* amplification of PDGFRA sequences, for instance the PDGFRA gene or the 5' or 3' flanking region thereof.

In some embodiments, kits may also include the reagents necessary to carry out nucleotide amplification reactions, including, for instance, DNA sample preparation reagents, appropriate

buffers *(e.g.,* polymerase buffer), salts (*e.g.*, magnesium chloride), and deoxyribonucleotides (dNTPs).

Kits may in addition include either labeled or unlabeled oligonucleotide probes for use in detection of PDGFRA mutation(s). In certain embodiments, these probes will be specific for a potential mutation that may be present in the target amplified sequences. The appropriate sequences for such a probe will be any sequence that includes one or more of the identified polymorphic sites, particularly nucleotide positions that overlap with the variants shown in Table 1 or Table 3, such that the sequence the probe is complementary to a polymorphic site and the surrounding PDGFRA sequence.

It may also be advantageous to provide in the kit one or more control sequences for use in the amplification reactions. The design of appropriate positive control sequences is well known to one of ordinary skill in the appropriate art.

B. Kits for Detection of PDGFRA mRNA Expression

Kits similar to those disclosed above for the detection of PDGFRA mutations directly can be used to detect PDGFRA mRNA expression, such as over- or under-expression. Such kits include an appropriate amount of one or more oligonucleotide primers for use in, for instance, reverse transcription PCR reactions, similarly to those provided above with art-obvious modifications for use with RNA amplification.

In some embodiments, kits for detection of altered expression of PDGFRA mRNA may also include some or all of the reagents necessary to carry out RT-PCR *in vitro* amplification reactions, including, for instance, RNA sample preparation reagents (including *e.g.*, an RNase inhibitor), appropriate buffers (*e.g.*, polymerase buffer), salts (*e.g.*, magnesium chloride), and deoxyribonucleotides (dNTPs). Written instructions may also be included.

Such kits may in addition include either labeled or unlabeled oligonucleotide probes for use in detection of the *in vitro* amplified target sequences. The appropriate sequences for such a probe will be any sequence that falls between the annealing sites of the two provided oligonucleotide primers, such that the sequence the probe is complementary to is amplified during the PCR reaction. In certain embodiments, these probes will be specific for a potential mutation that may be present in the target amplified sequences, for instance specific for the D842V or V561D point mutation, the ER561-562 in frame insertion, or the DIMH842-845, HDSN845-848P, RD841-842KI, RVIES560-564, or SPDGHE566-571R in-frame deletion, or another mutation identified in PDGFRA.

It may also be advantageous to provide in the kit one or more control sequences for use in the RT-PCR reactions. The design of appropriate positive control sequences is well known to one of ordinary skill in the appropriate art.

Alternatively, kits may be provided with the necessary reagents to carry out quantitative or semi-quantitative Northern analysis of PDGFRA mRNA. Such kits include, for instance, at least one PDGFRA-specific oligonucleotide for use as a probe. This oligonucleotide may be labeled in any conventional way, including with a selected radioactive isotope, enzyme substrate, co-factor, ligand,

chemiluminescent or fluorescent agent, hapten, or enzyme. In certain embodiments, such probes will be specific for a potential mutation that may be present in the target amplified sequence, such as the mutations disclosed herein.

C. Kits For Detection of PDGFRA Protein Expression

Kits for the detection of PDGFRA protein expression (such as over- or under-expression) are also encompassed. Such kits may include at least one target protein specific binding agent (*e.g.*, a polyclonal or monoclonal antibody or antibody fragment that specifically recognizes the PDGFRA protein) and may include at least one control (such as a determined amount of PDGFRA protein, or a sample containing a determined amount of PDGFRA protein). The PDGFRA-protein specific binding agent and control may be contained in separate containers. Likewise, kits for detection of activated PDGFRA may include at least one target protein binding agent (*e.g.*, a polyclonal or monoclonal antibody or antibody fragment) that specifically recognizes the PDGR-A protein only when PDGFRA is expressed in activated manner. These kits include, but are not limited to, those in which the PDGFRA binding agent recognizes, and binds specifically with, epitopes in which one or more tyrosine residues are phosphorylated. Kits for detection of activated/phosphorylated PDGFRA might include at least two controls, including a positive control with tyrosine phosphorylated PDGFRA and a negative control lacking tyrosine phosphorylated PDGFRA. The positive controls may include lysates or paraffin sections from cells and tissues expressing mutant (activated) PDGFRA, or expressing native PDGFRA that has been activated by exposure of the cells to PDGF-A. The negative controls may include lysates or paraffin sections from cells and tissues expressing non-activated PDGFRA, e.g. tissues expressing non-mutant PDGFRA, and without exposure to PDGF-A.

The PDGFRA protein expression detection kits may also include a means for detecting PDGFRA:binding agent complexes, for instance the agent may be detectably labeled. If the detectable agent is not labeled, it may be detected by second antibodies or protein A for example, which may also be provided in some kits in one or more separate containers. Such techniques are well known.

Additional components in specific kits may include instructions for carrying out the assay. Instructions will allow the tester to determine whether PDGFRA expression levels are elevated. Reaction vessels and auxiliary reagents such as chromogens, buffers, enzymes, etc. may also be included in the kits.

D. Kits for Detection of Homozygous versus Heterozygous Allelism

Also provided are kits that allow differentiation between individuals who are homozygous versus heterozygous for the D842V or V561D point mutations, the ER561-562 in frame insertion, or the DIMH842-845, HDSN845-848P, RD841-842KI, RVIES560-564, or SPDGHE566-571R in-frame deletion mutations of PDGFRA. Such kits provide the materials necessary to perform oligonucleotide ligation assays (OLA), as described at Nickerson *et al.* (*Proc. Natl. Acad. Sci. USA*

87:8923-8927, 1990). In specific embodiments, these kits contain one or more microtiter plate assays, designed to detect mutation(s) in the PDGFRA sequence of a subject, as described herein.

Additional components in some of these kits may include instructions for carrying out the assay. Instructions will allow the tester to determine whether a PDGFRA allele is homozygous or heterozygous. Reaction vessels and auxiliary reagents such as chromogens, buffers, enzymes, etc. may also be included in the kits.

It may also be advantageous to provide in the kit one or more control sequences for use in the OLA reactions. The design of appropriate positive control sequences is well known to one of ordinary skill in the appropriate art.

EXAMPLE 14: PDGFRA Knockout and Overexpression Transgenic Animals

Mutant organisms that under-express or over-express PDGFRA protein are useful for research. Such mutants allow insight into the physiological and/or pathological role of PDGFRA in a healthy and/or pathological organism. These mutants are "genetically engineered," meaning that information in the form of nucleotides has been transferred into the mutant's genome at a location, or in a combination, in which it would not normally exist. Nucleotides transferred in this way are said to be "non-native." For example, a non-PDGFRA promoter inserted upstream of a native PDGFRA encoding sequence would be non-native. An extra copy of a PDGFRA gene on a plasmid, transformed into a cell, would be non-native.

Mutants may be, for example, produced from mammals, such as mice, that either over-express PDGFRA or under-express PDGFRA, or that do not express PDGFRA at all. Over-expression mutants are made by increasing the number of PDGFRA genes in the organism, or by introducing an PDGFRA gene into the organism under the control of a constitutive or inducible or viral promoter such as the mouse mammary tumor virus (MMTV) promoter or the whey acidic protein (WAP) promoter or the metallothionein promoter. Mutants that under-express PDGFRA may be made by using an inducible or repressible promoter, or by deleting the PDGFRA gene, or by destroying or limiting the function of the PDGFRA gene, for instance by disrupting the gene by transposon insertion.

Antisense genes may be engineered into the organism, under a constitutive or inducible promoter, to decrease or prevent PDGFRA expression, as discussed above in Example 11.

A gene is "functionally deleted" when genetic engineering has been used to negate or reduce gene expression to negligible levels. When a mutant is referred to in this application as having the PDGFRA gene altered or functionally deleted, this refers to the PDGFRA gene and to any ortholog of this gene. When a mutant is referred to as having "more than the normal copy number" of a gene, this means that it has more than the usual number of genes found in the wild-type organism, *e.g.*, in the diploid mouse or human.

A mutant mouse over-expressing PDGFRA may be made by constructing a plasmid having a PDGFRA encoding sequence driven by a promoter, such as the mouse mammary tumor virus

(MMTV) promoter or the whey acidic protein (WAP) promoter. This plasmid may be introduced into mouse oocytes by microinjection. The oocytes are implanted into pseudopregnant females, and the litters are assayed for insertion of the transgene. Multiple strains containing the transgene are then available for study.

WAP is quite specific for mammary gland expression during lactation, and MMTV is expressed in a variety of tissues including mammary gland, salivary gland and lymphoid tissues. Many other promoters might be used to achieve various patterns of expression, *e.g.*, the metallothionein promoter.

An inducible system may be created in which the subject expression construct is driven by a promoter regulated by an agent that can be fed to the mouse, such as tetracycline. Such techniques are well known in the art.

A mutant knockout animal (*e.g.*, mouse) from which a PDGFRA gene is deleted can be made by removing all or some of the coding regions of the PDGFRA gene from embryonic stem cells. The methods of creating deletion mutations by using a targeting vector have been described (Thomas and Capecch, *Cell* 51:503-512, 1987).

Engineered PDGFRA knockout animals are known. See, for instance, Bostrom *et al.*, *Dev. Dyn.*, 223:155-162, 2002; Fruttiger *et al.*, *Development*, 126:457-467, 1999; Hellstrom *et al.*, *J. Cell Biol.*, 153:543-553, 2001; Kaminski *et al.*, *Blood*, 97:1990-1998, 2001; Karlsson *et al.*, *Development*, 127:3457-3466, 2000. In addition, *Patch* mutant mice have a congenital chromosomal deletion that includes the PDGFR-α gene locus.

Example 15: Knock-in Organisms

In addition to knock-out systems, it is also beneficial to generate "knock-ins" that have lost expression of the wildtype protein but have gained expression of a different, usually mutant form of the same protein. By way of example, the activating mutant PDGFRA mutant proteins provided herein (*e.g.*, as shown in SEQ ID NO: 4, 6, 8, 10, 12, 21, 23, 25, and 27) can be expressed in a knockout background, such as the *Patch* mutant mice, in order to provide model systems for studying the effects of these mutants. In particular embodiments, the resultant knock-in organisms provide systems for studying neoplasia.

Those of ordinary skill in the relevant art know methods of producing knock-in organisms. See, for instance, Rane *et al.* (*Mol. Cell Biol.*, 22: 644-656, 2002); Sotillo *et al.* (*EMBO J.*, 20: 6637-6647, 2001); Luo *et al.* (*Oncogene*, 20: 320-328, 2001); Tomasson *et al.* (*Blood*, 93: 1707-1714, 1999); Voncken *et al.* (, 86: 4603-4611, 1995); Andrae *et al.* (*Mech. Dev.*, 107: 181-185, 2001); Reinertsen *et al.* (*Gene Expr.*, 6: 301-314, 1997); Huang *et al.* (*Mol. Med.*, 5: 129-137, 1999); Reichert *et al.* (*Blood*, 97: 1399-1403, 2001); and Huettner *et al.* (*Nat. Genet.*, 24: 57-60, 2000), by way of example.

Example 16: Demonstration of PDGFRA Fusion Oncoproteins in Human Leukemias.

The PDGFRA activating genomic mutations disclosed herein involve intragenic point mutations or deletions. These models of genomic PDGFRA mutation can readily be extended to different mechanisms of activation, *e.g.* as might result from chromosomal rearrangement in which the promoter and 5' end of an ectopic gene are fused to the 3' end – including the kinase domain – of PDGFRA. The principle of receptor tyrosine kinase activation, in which cytogenetic rearrangement produces a gene fusion, has been established for several kinase proteins, including FGFR1, FGFR3, NTRK3, and ALK, and have been reported recently for PDGFRA, in two patients with chronic myelogenous leukemia, in which PDGFRA was fused with the BCR gene. In the PDGFRA context, the applicants have identified four leukemias in which cytogenetic banding analyses reveal translocation breakpoints in the PDGFRA gene (chromosome band 4q12) region, and in which – based on cytogenetic correlates - the putative PDGFRA fusion gene is not expected to be BCR. Therefore, these leukemias may contain novel forms of PDGFRA fusion oncogenes. FISH analyses will be performed to determine whether any of these translocations targets PDGFRA, in which case the translocation partner gene will be identified by rapid amplification of cDNA ends, and the activating nature of the PDGFRA fusion will be determine by expressing the PDGFRA fusion gene in cell types such as Ba/F3 and CHO.

Example 17: Additional PDGFRA Activating Mutations in Gastrointestinal Stromal Tumors

Using methods essentially as described in Example 1, three additional PDGFRA activating mutations were identified in GISTs. These mutations are as shown in Table 3.

Table 3

Genotype	DNA sequence (top line) Translation (bottom line)
PDGFRA Wild type (SEQ ID NOs: 1 and 2)	2906* GGCCTGGCCAGAGACATCATGCATGATTCGAACTATGTG 838 G L A R D I M H D S N Y V
PDGFRA Deletion RD841-842KI (SEQ ID NOs: 24 and 25)	2906 GGCCTGGCCAAAATCATCATGCATGATTCGAACTATGTG 838 G L A K I I M H D S N Y V
PDGFRA Wild type	2060 GAAATTCGCTGGAGGGTCATTGAATCAATCAGCCCGGAT 556 E I R W R V I E S I S P D
V561D (SEQ ID NOs: 20 and 21)	2060 GAAATTCGCTGGAGGGACATTGAATCAATCAGCCCGGAT 556 E I R W R D I E S I S P D
PDGFRA Deletion RVIES560-564 (SEQ ID NOs: 22 and 23)	2060 GAAATTCGCTGG----------ATCAGCCCGGAT 556 E I R W - - - - - I S P D

*Numbering as in SEQ ID NO: 1 and SEQ ID NO: 2.

After taking into account these three additional mutations, and additional instances of other identified mutations, the total number of each of the identified activating mutations was as shown in Table 4 and Table 5.

Table 4. Summary of *PDGFRA* mutations in *KIT*-WT GISTs.

PDGFRA Region	Mutation	#GISTs
Activation Loop (exon 18)	D842V	15
	Del DIMH	4
	Del HDSN845-848P	1
	Del RD841-842KI	1
Juxtamembrane (exon 12)	V561D	1
	Ins ER561-562	1
	Del RVIES560-564	1
	Del SPDGHE566-571R	1

Table 5.

Mutation	Cases (% total)
D842V	15 (21.7%)
Exon 18 Deletion	6 (8.7%)
Exon 12 Insertion/Deletion/PM	4 (5.8%)
No mutation	44 (63.7%)
Total	**69 (100.0%)**

The nucleic acid sequences of all of the identified activating PDGFRA mutations were aligned to produce the consensus sequence shown in SEQ ID NO: 26; the numbering in the consensus sequence aligns with that in the wildtype PDGFRA nucleic acid sequence (SEQ ID NO: 1). In the consensus sequence, the insertion identified in variant PDGFRA Insertion ER561-562 is indicated in a miscellaneous features field in the Sequence Listing. As emphasized and clearly illustrated in the consensus sequence, clusters of activating mutations in the PDGFRA nucleic acid sequence are found in positions 2072 to 2107 and 2916 to 2937, though it is noted that positions 2087, 2088, and 2089 appear to be invariable at least in the current studies.

Example 18: Additional Characterization of PDGFRA Activating Mutations in GISTs

Materials and Methods

Reagents

Antibodies used for immunoblotting were to phosphotyrosine (Santa Cruz PY99), actin (Sigma 1PKCA4), KIT (Dako A4502), PDGFRA (Santa Cruz sc-338), phosphoPDGFRA Y754 (Santa Cruz sc-12911), MAPK (Zymed 61-7400), phosphoMAPK Thr202/Thr204 (Cell Signaling 9106), AKT (Cell Signaling 9272), phosphoAKT S473 (Cell Signaling 9271S), STAT1 (Zymed ST1-3D4), phosphoSTAT1 Y701, (Zymed ST1P-11A5), STAT3 (Zymed 13-7000), phosphoSTAT3 Y705 (Cell Signaling 9131), STAT5 (Zymed ST5-8F7), and phosphoSTAT5 Y694 (Zymed ST5P-4A9). Antibodies to phosphorylated kinases were validated as phosphospecific by evaluation of phosphatase treated cell lysates, and by evaluation of lysates from GIST cells treated with kinase inhibitors.

Cytogenetic Analyses

Tumor specimens were chopped with scalpel blades, disaggregated enzymatically, and seeded into T25 flasks. The monolayer cultures were expanded for two-to-five days prior to metaphase cell harvesting with Colcemid. Tissue culture, metaphase harvesting, metaphase slide making, and Giemsa-trypsin banding were performed as described previously (Fletcher *et al.*, *N. Engl. J. Med.* 324, 436, 1991).

Cloning, Expression and Characterization of PDGFRA mutant cDNAs

PDGFRA mutations were cloned by site-directed mutagenesis of the wild type *PDGFRA* cDNA. CHO cells were transiently transfected with expression vectors encoding for mutant or wild-type *PDGFRA* cDNA. Transfected cells were serum starved overnight and stimulated with vehicle or 100 ng/ml recombinant human PDGF-AA for 10 minutes before harvesting cells and preparing whole cell lysates for immunoblotting. The membranes were sequentially immunoblotted with antiserum against phosphorylated tyrosines (PY20 Transduction Laboratories) or total PDGFRA (Santa Cruz sc-338).

Results and Discussion

The biochemical consequences of somatic PDGFRA mutations were studied by transient expression of wild-type and mutant *PDGFRA* cDNA constructs in Chinese hamster ovary (CHO) cells. Baseline tyrosine phosphorylation was weak for non-mutant PDGFRA, and was substantially increased by ligand stimulation (Figure 8). By contrast, baseline tyrosine phosphorylation was strong in all five of the tested PDGFRA mutants, and was not increased by ligand stimulation (Figure 8).

Next the signal transduction pathways activated in PDGFRA-mutant versus KIT-mutant GISTs were compared. The PDGFRA-mutant GISTs showed uniform activation of signaling intermediates AKT, MAPK, STAT1, and STAT3, which are also activated in most KIT-mutant GISTs (Figure 9). The PDGFRA-mutant GISTs lacked expression of phosphoSTAT5, despite strong expression of total STAT5, which is also typical of KIT-mutant GISTs. The cytogenetic profiles of four PDGFRA-mutant GISTs and 52 KIT-mutant GISTs were also compared. KIT mutations are early events in GIST tumorigenesis, whereas cytogenetic aberrations occur later in disease progression (Heinrich *et al.*, *Hum. Pathol.* 33, 484, 2002). Most of these GISTs – irrespective of *PDGFRA* or *KIT* mutation – featured noncomplex karyotypes with deletions of chromosome 1p, and with monosomies of chromosomes 14 and 22. Hence, these results suggest that the mechanisms of cytogenetic progression and oncoprotein-driven signal transduction are similar in GISTs expressing oncogenic forms of PDGFRA and KIT.

Activating mutations of KIT or PDGFRA appear to be mutually exclusive oncogenic events in GISTs, and these mutations have similar biological consequences. The data presented also highlight a crucial role for PDGFRA in the pathogenesis of a solid tumor. Notably, a translocation involving the *BCR* and *PDGFRA* genes has been described in BCR-ABL negative chronic myelogenous leukemia, and is predicted to result in dimerization and kinase activation of the fusion

protein (Baxter *et al.*, *Hum. Mol. Genet.* 11, 1391, 2002). PDGFRA is widely expressed in human tissues, so it will be important to determine whether PDGFRA mutations play a role in other human malignancies. Such tumors could be sensitive to Gleevec and other small molecule drugs that inhibit PDGFRA kinase activity (Buchdunger *et al.*, *J. Pharmacol. Exp. Ther.* 295, 139, 2000; Lokker *et al.*, *Cancer Res.* 62, 3729, 2002; Sun *et al.*, *J. Med. Chem.* 43, 2655, 2000).

This disclosure provides tyrosine kinase protein and nucleic acid variants, particularly PDGFRA variants, which are activating forms of these molecules and are linked to neoplasms and/or the development or progression of cancer. The disclosure further provides methods of diagnosis and prognosis, using these molecules and fragments thereof, and kits for employing these methods and compositions. It will be apparent that the precise details of the compositions and methods described may be varied or modified without departing from the spirit of the described invention. We claim all such modifications and variations that fall within the scope and spirit of the claims below.